

QY 3341 GTTACCTGGACATTCACCACTCAGGATGATTGCCCCAGTGAA 2400
 2341 GTTACCTGGACATTCACCACTCAGGATGATTGCCCCAGTGAA 2400
 Db 2401 CATACCAAGCTGTGAATTGGAAAAAATAATTGTTGTCCTGGGAA 2400
 QY 2401 CATACCAAGCTGTGAATTGGAAAAAATAATTGTTGTCCTGGGAA 2400
 Db 2401 ACTTGTATGGCCACAAACTAAACCGTGTGGGGAGCTGG 2400
 QY 2401 ACTTGTATGGCCACAAACTAAACCGTGTGGGGAGCTGG 2400
 Db 2401 ACTTGTATGGCCACAAACTAAACCGTGTGGGGAGCTGG 2400
 QY 2521 CAGATTTCACGGTACATTAACTCCCAAACTACCGGAAATTACCCAGC 2580
 Db 2521 CAGATTTCACGGTACATTAACTCCCAAACTACCGGAAATTACCCAGC 2580
 QY 2581 AGTGTACGTGGGACCATAACCCCAAGCGGGCATCTGATGTGGTCTGAGA 2640
 Db 2581 AGTGTACGTGGGACCATAACCCCAAGCGGGCATCTGATGTGGTCTGAGA 2640
 QY 2641 TCTTCCCTGCCATAGGGACGACTGTGGGACATATGGTGTGGGAAACCTCTCAT 2700
 Db 2641 TCTTCCCTGCCATAGGGACGACTGTGGGACATATGGTGTGGGAAACCTCTCAT 2700
 QY 2701 CCAATTCTGACAACTATGAAACCTGCCCACCTACGGCTTCACCT 2760
 Db 2701 CCAATTCTGACAACTATGAAACCTGCCCACCTACGGCTTCACCT 2760
 QY 2761 CCAGGTCAAAGAAAGCTGTGGATTCAAGTCAGTCAAGGGAAACAGGGCTACAGGGT 2820
 Db 2761 CCAGGTCAAAGAAAGCTGTGGATTCAAGTCAGTCAAGGGAAACAGGGCTACAGGGT 2820
 QY 2821 TCCAGGTCCTAACAGTGTACATATGATGAGGACTACAGGAACTCATGGAGACATAGTC 2880
 Db 2821 TCCAGGTCCTAACAGTGTACATATGATGAGGACTACAGGAACTCATGGAGACATAGTC 2880
 QY 2881 SAGATGCAGGGCTATGCAATTGAGAACCTCAGGAAATCTTCTAGGATAGAACCTTA 2940
 Db 2881 SAGATGCAGGGCTATGCAATTGAGAACCTCAGGAAATCTTCTAGGATAGAACCTTA 2940
 QY 2941 TCAAGGTCTCTTTGATGTCCTGGCCCATCAGTCAACGCCAGG 3000
 Db 2941 TCAAGGTCTCTTTGATGTCCTGGCCCATCAGTCAACGCCAGG 3000
 QY 3001 AGTCCCGAGAGATGTTCCAAGATGTCCTGGCCCATCAGTCAACGCCAGG 3060
 Db 3001 AGTCCCGAGAGATGTTCCAAGATGTCCTGGCCCATCAGTCAACGCCAGG 3060
 QY 3061 TTGTTGAGACCTACAAAAGTCAGCTGCTGGCCACTCATACAAATGTTCTGTATAG 3120
 Db 3061 TTGTTGAGACCTACAAAAGTCAGCTGCTGGCCACTCATACAAATGTTCTGTATAG 3120
 QY 3121 GGTGGTGGAGAGGTGTGTTCTGGATGTAGCAAGTCGGGTATGTCGGCTC 3180
 Db 3121 GGTGGTGGAGAGGTGTGTTCTGGATGTAGCAAGTCGGGTATGTCGGCTC 3180
 QY 3181 CGGTATCAGTGAACCTCAATTGACTCAATTAGATAATACAGATAATTGGTAATT 3240
 Db 3181 CGGTATCAGTGAACCTCAATTGACTCAATTAGATAATACAGATAATTGGTAATT 3240
 QY 3241 GAACCTGCTTCCTCCACGATGTCGGGAGATGTCGGTCTGGATGAGAATGCTTGAACGCCAT 3300
 Db 3241 GAACCTGCTTCCTCCACGATGTCGGGAGATGTCGGTCTGGATGAGAATGCTTGAACGCCAT 3300
 QY 3301 CAGCTTCTCACTGCTGGGGAGATGTCGGTCTGGATGAGAATGCTTGAACGCCAT 3360
 Db 3301 CAGCTTCTCACTGCTGGGGAGATGTCGGTCTGGATGAGAATGCTTGAACGCCAT 3360
 QY 3361 TGGTAGCCATTGGAGACTAACCTGCTCCCTGGGGTCTGGATGAGAATGCTTGAACGCCAT 3420
 Db 3361 TGGTAGCCATTGGAGACTAACCTGCTCCCTGGGGTCTGGATGAGAATGCTTGAACGCCAT 3420
 QY 3421 TGATGTGAAAGGGGACAGATAAGCTGCTTATISTGAAACTCTAGG 3480
 Db 3421 TGATGTGAAAGGGGACAGATAAGCTGCTTATISTGAAACTCTAGG 3480
 QY 3481 CGGGCCCTCTCTAGGGACCCCTGCACTGCTGGGGCTGACAGGGAAGAGCG 3540
 Db 3481 CGGGCCCTCTCTAGGGACCCCTGCACTGCTGGGGCTGACAGGGAAGAGCG 3540
 QY 3544 CAAGGGGAGGGAGAGAACCTCTGAGGGCTCTGACCTGGGG 3600
 Db 3544 CAAGGGGAGGGAGAGAACCTCTGAGGGCTCTGACCTGGGG 3600
 QY 3601 ACTAGTGTCTCCACGCTTCCAGGCTCTGAGGGCTCTGACCTGGGGACTG 3660
 Db 3601 ACTAGTGTCTCCACGCTTCCAGGCTCTGAGGGCTCTGACCTGGGGACTG 3660
 QY 3661 AGCTCTAAGAGGTGTCTGTAAGTGTAGTCCAGGAAATTAGA 3720
 Db 3661 AGCTCTAAGAGGTGTCTGTAAGTGTAGTCCAGGAAATTAGA 3720
 QY 3721 AGCACTCTGGAGACAT 3737
 Db 3721 AGCACTCTGGAGACAT 3737
 QY 3721 AGCACTCTGGAGACAT 3737

RESULT 2
 US-13-177-293-44
 Sequence 44, Application US/13-177-293-44
 Publication No. US20030124128A1
 GENERAL INFORMATION:
 APPLICANT: Lillie, James
 APPLICANT: Glatt, Karen
 APPLICANT: Zhao, Xumei
 APPLICANT: Gharavarpur, Manjula
 APPLICANT: Kamatkar, Shubhangi
 APPLICANT: Verzeni, Maureen
 APPLICANT: Wyer, Vic
 APPLICANT: Wang, Youzhen
 APPLICANT: Xu, Yongyac
 APPLICANT: Hoersch, Sebastian
 APPLICANT: Monahan, John
 APPLICANT: Meyers, Racine E.
 APPLICANT: Bast, Christopher
 APPLICANT: Horroabay, Gabriel N.
 APPLICANT: Puszta, Lajos
 APPLICANT: Meric, Funda
 APPLICANT: Sahin, Aysegul
 APPLICANT: Milis, Gordon B.
 TITLE OF INVENTION: COMPOSITIONS, KITS AND METHODS FOR IDENTIFICATION, ASSESSMENT AND THERAPY OF BREAST CANCER
 FILE REFERENCE: MRI-038
 CURRENT APPLICATION NUMBER: US/13-177-293-44
 CURRENT FILING DATE: 2003-02-26
 PRIOR APPLICATION NUMBER: US 60/124128
 PRIOR FILING DATE: 2001-06-21
 PRIOR APPLICATION NUMBER: US 60/1321572
 PRIOR FILING DATE: 2001-06-27
 PRIOR APPLICATION NUMBER: US 60/1326501
 PRIOR FILING DATE: 2001-07-08
 PRIOR APPLICATION NUMBER: US 60/1325022
 PRIOR FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US 60/1362595
 PRIOR FILING DATE: 2002-03-05
 PRIOR APPLICATION NUMBER: US 60/1325022
 NUMBER OF SEQ ID NCS: 506
 SOFTWARE: FastSEQ for Windows version 4.0
 SEQ ID NO: 64
 LENGTH: 3737
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-13-177-293-44

Chery Marc,
 Score 1777, CR 141, Iterat 1737.

Best Local Similarity	100.0%	Pred. No.	0	Mismatches	0	Indels	0	Gaps	3
Matches	3737	Conservative	0						
1	GGCGTCCGGCAGACCTCCCCGGCGCGCACCGCCGCCAACCGCCGCACTCCGGCCCTCTGC	60	D _b						
1	GGGTCCCGAACCCTCCCGCGCGCACCGCCGCCAACCGCCGCACTCCGGCCCTCTGC	60	2Y						
61	CGCCAACGGCATGCCATCCATGGGTCGCGGCGAACCTCCGGCGAACCTCCGGCGCTGG	120	D _b						
61	CGCCAACGGCATGCCATCCATGGGTCGCGGCGAACCTCCGGCGAACCTCCGGCGCTGG	120	Q _y						
121	CGGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG	180	D _b						
121	CGGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG	180	Q _y						
181	CGGTGCGGGCGCCGGCGAACGGGAACTGAGTGAGTGAGTGAGTGAGTGAGTGAGTG	240	D _b						
181	CGGTGCGGGCGCCGGCGAACGGGAACTGAGTGAGTGAGTGAGTGAGTGAGTGAGTG	240	Q _y						
241	AATGACTGGCATGGCACGCCCTGTCTGAGAACCCACTCTTAAGTGCTCTCGCA	300	D _b						
241	AATGACTGGCATGGCACGCCCTGTCTGAGAACCCACTCTTAAGTGCTCTCGCA	300	25						
301	AAGCTGGTTACCAAGGGAAGGGAAGGCACTGGTGTGAGCATGGTAAATGGAAATAGC	360	D _b						
301	AAGCTGGTTACCAAGGGAAGGGAAGGCACTGGTGTGAGCATGGTAAATGGAAATAGC	360	Q _y						
361	TCAATGGGGCTGTGTCATGACTCTGTTGAAATTCCGGCAATTATGTTGCACTGTG	420	D _b						
361	TCAATGGGGCTGTGTCATGACTCTGTTGAAATTCCGGCAATTATGTTGCACTGTG	420	Q _y						
421	TGATGGCTTCACTGGCTCATGACGGCTCATATTGCTGTTGAGCTGAGTGCTGG	480	D _b						
421	TGATGGCTTCACTGGCTCATGACGGCTCATATTGCTGTTGAGCTGAGTGCTGG	480	Q _y						
481	AGACAAATGGGGCTGGCTGGAGCATACCTGTCAACGGTCAATGGGAGCTGGCT	540	D _b						
481	AGACAAATGGGGCTGGCTGGAGCATACCTGTCAACGGTCAATGGGAGCTGGCT	540	Q _y						
541	GCAGGGAGGGCTTTCTGAGTGCACATAGGACACCTGCACTACCGCTGCACTCGA	600	D _b						
541	GCAGGGAGGGCTTTCTGAGTGCACATAGGACACCTGCACTACCGCTGCACTCGA	600	Q _y						
601	GCTCTGAGGTGCTGAGTAATAGGATCACTGGCTGTAGTCACATCTGAGGAAAGGG	660	D _b						
601	GCTCTGAGGTGCTGAGTAATAGGATCACTGGCTGTAGTCACATCTGAGGAAAGGG	660	Q _y						
661	GCAGCTGGCTGGCTGGAGCTGGTGTGAGTGGCTGGAGCTGGCTGGAGCTGGAG	720	D _b						
661	GCAGCTGGCTGGCTGGAGCTGGTGTGAGTGGCTGGAGCTGGCTGGAGCTGGAG	720	Q _y						
721	TCTTGACCTGTAACATGGAACTGGTGTGAGTCACTGGTGTGAGTCACTGGTGT	780	D _b						
721	TCTTGACCTGTAACATGGAACTGGTGTGAGTCACTGGTGTGAGTCACTGGTGT	780	Q _y						
781	GCCCAGACTGCACTGGTGTGAGTCACTGGTGTGAGTCACTGGTGTGAGTCACTGG	840	D _b						
781	GCCCAGACTGCACTGGTGTGAGTCACTGGTGTGAGTCACTGGTGTGAGTCACTGG	840	Q _y						
841	AGCGAACAGAACACTGCTGGTGTGAGTCACTGGTGTGAGTCACTGGTGTGAGT	900	D _b						
841	AGCGAACAGAACACTGCTGGTGTGAGTCACTGGTGTGAGTCACTGGTGTGAGT	900	Q _y						
901	ATAACGGGTGAAACGGGGCTGTGCTGTCATGGAACTGGTGTGAGTCACTGGTGT	960	D _b						
901	ATAACGGGTGAAACGGGGCTGTGCTGTCATGGAACTGGTGTGAGTCACTGGTGT	960	Q _y						
.961	ACCGAACCTGTAAGGATACTTCGACGGTGTGACTGAGTGTGCTGTTGATTCATC	1020	D _b						
.961	ACCGAACCTGTAAGGATACTTCGACGGTGTGACTGAGTGTGCTGTTGATTCATC	1020	Q _y						
961	ACCCACCTGTAAGGATACTTCGACGGTGTGACTGAGTGTGCTGTTGATTCATC	1020	D _b						
961	ACCCACCTGTAAGGATACTTCGACGGTGTGACTGAGTGTGCTGTTGATTCATC	1020	Q _y						
1021	TCCAGTGGTGGAAAGAACATGTGAAATGGGAGCTGTTGGAAATGGGAGCTGTT	1080	D _b						
1021	TCCAGTGGTGGAAAGAACATGTGAAATGGGAGCTGTTGGAAATGGGAGCTGTT	1080	Q _y						

Db	3381	ATGTTATCGTGTGCCAGTGGGAAACATAACCCAGCTGATTGAAAAAATATGTGGTTTC 2449
Qy	2438	TTGCCCAAGGAATAACTACGACTGACTTTGATGGCTCACAACATAACCCAGTGTAAA 2497
Db	2441	TTGCCCAAGGAATAACTACGACTGACTTTGATGGCTCACAACATAACCCAGTGTAAA 2500
Qy	2498	CAGAAGATGTGGAGCTGGAGCTGGGAAATTTCATGGTACATGATACCCAAATACCC 2557
Db	2501	CAGAAGATGTGGAGCTGGGAAATTTCATGGTACATGATACCCAAATACCC 2560
Qy	2558	AGGCAATTACCCAGCACAACACGGGTTAACCTGGACCATCAACCCACCCCAAGGCCG 26-7
Db	2561	AGGCAATTACCCAGCACAACACGGGTTAACCTGGACCATCAACCCACCCCAAGGCCG 26-23
Qy	2618	CATCCCTATCGTGTCCCTGTGATCTCTCCATAGAGAACGACTGTTGGACTATCT 2677
Db	2621	CATCCCTATCGTGTCCCTGTGATCTCTCCATAGAGAACGACTGTTGGACTATCT 2680
Qy	2678	GCTGATTCGGAAAACCCTCTTACATCCATTCTGTGACACATAGAACCTGGCAACTTA 2737
Db	2681	GCTGATTCGGAAAACCCTCTTACATCCATTCTGTGACACATAGAACCTGGCAACTTA 2740
Cy	2738	CGAACGCCCATCGCCCTAACCTCCAGCTCAAAGAACGCTGGGATCTAGTCAAGTCAA 2797
Db	2741	CGAACGCCCATCGCCCTAACCTCCAGCTCAAAGAACGCTGGGATCTAGTCAAGTCAA 2800
Qy	2798	TGAAGGGAAACACGGCTAGAGGGTCTCAGGTCATAGTGCATATGATGAGGATACCA 2857
Db	2801	TGAAGGGAAACACGGCTAGAGGGTCTCAGGTCATAGTGCATATGATGAGGATACCA 2860
Qy	2858	GAAACTCATGAGAACTAGTCGATGGGGCTATCATCTGAGAACCTATGAGA 2917
Db	2861	GAAACTCATGAGAACTAGTCGATGGGGCTATCATCTGAGAACCTATGAGA 2920
Qy	2918	AATACTTAAGGATAAGAACTTATCAAGGCTCTGTTGATCTCCGGCCATCCCAGAA 2977
Db	2921	AATACTTAAGGATAAGAACTTATCAAGGCTCTGTTGATCTCCGGCCATCCCAGAA 2980
Qy	2978	CTATTCAACTACAGCCAGAGTCCGGAGATGTTCAAGAGTCGTCTCATCGATT 3037
Db	2981	CTATTCAACTACAGCCAGAGTCCGGAGATGTTCAAGAGTCGTCTCATCGATT 3040
Qy	3038	GCTACGGTCAAAAGTGGCAGGGTTTGAGACCTTACAATGACTGAGCCACGTGCCAC 3097
Db	3041	GCTACGGTCAAAAGTGGCAGGGTTTGAGACCTTACAATGACTGAGCCACGTGCCAC 3105
Qy	3098	TCAATACAAAATGTTCTGCTATAGGGTTGGGGCACAGAGCTGTCCTCTGGCTGTCA 3157
Db	3101	TCAATACAAAATGTTCTGCTATAGGGTTGGGGCACAGAGCTGTCCTCTGGCTGTCA 3160
Qy	3158	SCACAGTCGGTATTGTCGCTCCGGTATCAGTGTACATGCTTGTGAGTCAATTGTATAGA 3217
Db	3161	SCACAGTCGGTATTGTCGCTCCGGTATCAGTGTACATGCTTGTGAGTCAATTGTATAGA 3220
Qy	3218	TAATACAGATAATTGTTGTTAATGAACTGGTTCTCACTGCTGTTCTGGATATAGA 3277
Db	3221	TAATACAGATAATTGTTGTTAATGAACTGGTTCTCACTGCTGTTCTGGATATAGA 3280
Qy	3278	CTGAGAATGGCTTGTGACTGGCATCAGCTTCTCACTGCTGTTCTGGATATAGA 3337
Db	3281	CTGAGAATGGCTTGTGACTGGCATCAGCTTCTCACTGCTGTTCTGGATATAGA 3340
Qy	3338	TCACGGACTGGCTGAGCTGGATTTCATGGCTTGTGAGTCACTGCTGTTCTGG 3397
Db	3341	TCACGGACTGGCTGAGCTGGATTTCATGGCTTGTGAGTCACTGCTGTTCTGG 3400
Qy	3398	GTCCTATCCTCCTCAAGGAGTCGTGAGTGGAAAGGGGCCACAGATAAGCTGTTCT 3457
Db	3401	GTCCTACTCCTCCTCAAGGAGTCGTGAGTGGAAAGGGGCCACAGATAAGCTGTTCT 3460
Qy	3458	CTGAAACTTCAGTTCTCTCTAGCCGCCCTCTCAAGGGAGCCCTCTGACTCTGTG 35-7
Db	3461	CTGAAACTTCAGTTCTCTAGCCGCCCTCTCAAGGGAGCCCTCTGACTCTGTG 35-2

Db	401	AGGCAATTATCGTGGCACTGTGTTGATGGCTCATGTCGCTCATGACGGCTATAATTG	460
Qy	458	TCTTGATSTGGCAGACTGCTGGAGAACATACTGGCCCTGCCAGATACTCTGTGTCACGT	517
Db	461	TCTTGATGGGGAGAGTCGCTGGAGAACATACTGGCCCTGCCAGATACTCTGTGTCACGT	520
Qy	518	CATGGGGAGCTATGAGTGCTGTGGAGAACAGGGGGTTTCTGTAGTGACAATAGGACAC	577
Db	521	CATGGGGAGCTATGAGTGCTGTGGAGAACAGGGGGTTTCTGTAGTGACAATAGGACAC	580
Qy	578	CTGCATTCACCCTCGAAAGAGGGCTGTAGTGCTGAATTAAGATCAGGGCTGTAGTGCA	637
Db	581	CTGCATTCACCCTCGAAAGAGGGCTGTAGTGCTGAATTAAGATCAGGGCTGTAGTGCA	640
Qy	638	CATCTGAAAGGGCCCAGGGCAGCTGGCTGTAGTGCTGGGGCTGTAGTGCTGTAGTGCT	697
Db	641	CATCTGAAAGGGCCCAGGGCAGCTGGCTGTAGTGCTGGGGCTGTAGTGCTGTAGTGCT	700
Qy	698	GCCCCAAACCCAGAGAACATGCACTCTGACCTGTAACTGGGACGGCTGTAGTGCA	757
Db	701	GCCCCAAACCCAGAGAACATGCACTCTGACCTGTAACTGGGACGGCTGTAGTGCA	760
Qy	758	CTCTGTGACGATACAGCGATGGCCATCCACAGTACAGATGCA	817
Db	761	CTCTGTGACGATACAGCGATGGCCATCCACAGTACAGATGCA	820
Qy	818	CACAGATGGAGGAGCTGGCTTGGAGGAGGACAGTCTCTGACCTGTACCGAGGGCT	877
Db	821	CACAGATGGAGGAGCTGGCTTGGAGGAGGACAGTCTCTGACCTGTACCGAGGGCT	880
Qy	878	CACCACTAGTGGATGGGATAAACGGTGAACGGGGTGTCACTGAAACACTG	937
Db	881	CACCACTAGTGGATGGGATAAACGGTGAACGGGGTGTCACTGAAACACTG	940
Qy	939	TGCTGTGACATGGGGCTGTGAGGACCTGTAGGATACTGACAGG-GTTCACAT	997
Db	941	TGCTGTGACATGGGGCTGTGAGGACCTGTAGGATACTGACAGG-GTTCACAT	1000
Qy	998	CAGTGTCTGTGGATGGGCTACTCTGGATGGGAAACATGGTAAAGATATTGATA	1057
Db	1001	CAGTGTCTGTGGATGGGCTACTCTGGATGGGAAACATGGTAAAGATATTGATA	1060
Qy	1058	GTGCCAGACCCCAATGGGGCTGTGAGGATCTTCTGAAACATGGTAAAGATATTGATA	1117
Db	1061	GTGCCAGACCCCAATGGGGCTGTGAGGATCTTCTGAAACATGGTAAAGATATTGATA	1120
Qy	1118	CTGGGGCTGCAAGAAGGATTAAATTATTAACATGAGAACATGGTAAAGATATTGATA	1177
Db	1121	CTGGGGCTGCAAGAAGGATTAAATTATTAACATGAGAACATGGTAAAGATATTGATA	1180
Qy	1178	TGAGTGTCTTGGATGGACCTGACCCAGCTGACCTGGCAACCTGGCAATTC	1237
Db	1181	TGAGTGTCTTGGATGGACCTGACCCAGCTGACCTGGCAACCTGGCAATTC	1240
Qy	1238	TCTGCTGCAACCGGAGGTACACCCGTACCCACTGTGGGACACCAATGA	1297
Db	1241	TCTGCTGCAACCGGAGGTACACCCGTACCCACTGTGGGACACCAATGA	1300
Qy	1358	ATGCCASTGCCACCCCTGGGATCAAAGCTCCACTGGAAATAAAAAGACTGTGGAAGGTGA	1417
Db	1361	ATGCCASTGCCACCCCTGGGATCAAAGCTCCACTGGAAATAAAAAGACTGTGGAAGGTGA	1420
Qy	1418	GGGGCTCTGCCACACAGTGTGCTGCCCTGTCACCCGCTGTCACCTGGCTGTAGCTGG	1477
Db	1421	GGGGCTCTGCCACACAGTGTGCTGCCCTGTCACCCGCTGTCACCTGGCTGTAGCTGG	1480
Qy	1478	AGGAGACGGTGTCTCTGAGTGTCACTCTGCACTGAGTGTCACTGAGTGTCACTG	1537
Db	1481	AGGAGACGGTGTCTCTGAGTGTCACTCTGCACTGAGTGTCACTGAGTGTCACTG	1540
Qy	1538	CATCGAGACAGTGTAACTTGTAGCTAAATGAGGAAAGTGTGAAATGTGA	1597
Db	1541	CATCGAGACAGTGTAACTTGTAGCTAAATGAGGAAAGTGTGAAATGTGA	1600
Qy	1598	SCTGTTTCCGAGGCTTCGACGGACTACCGAGACAGCTAAAGAGAG	1657
Db	1601	GC:GTTCGAGGCTTCGACGGACTACCGAGACAGCTAAAGAGAG	1660
Qy	1658	CITCGCGTAGCTGAACTCTGACCTGAGCTGCTGCTGCG	1717
Db	1661	CITCGCGTAGCTGAACTCTGACCTGAGCTGCTGCTGCG	1720
Qy	1718	ACCAAGCACCCCTGGAGATGTTTACCTGTGAGCTTGTGAGCTGAACTAACCTACAA	1777
Db	1721	ACCAAGCACCCCTGGAGATGTTTACCTGTGAGCTTGTGAGCTGAACTAACCTACAA	1780
Qy	1778	GGAGGTGACAGCTCTTGAGCTGCTGATGTAAGGAGACAGAGGGCTCG	1837
Db	1781	GGAGGTGACAGCTCTTGAGCTGCTGATGTAAGGAGACAGAGGGCTCG	1840
Qy	1838	TAAAGGATCTGGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	1897
Db	1841	TAAAGGATCTGGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	1900
Qy	1898	AGGCCATGAACTCTGACGTGGCTTAAAGCTCCAGAACATCGAACGCCAGCA	1957
Db	1901	AGGCCATGAACTCTGACGTGGCTTAAAGCTCCAGAACATCGAACGCCAGCA	1960
Qy	1958	CTGTGGATGGGGCCGCGCTCATGAGAAACCATGTTGAGCTGCGGACCTA	2017
Db	1961	CTGTGGATGGGGCCGCGCTCATGAGAAACCATGTTGAGCTGCGGACCTA	2020
Qy	2018	TTATGTTGGGACGAGAACCTGCTGCAATTGCTGCAATTGCTGCAATTGCTGCA	2077
Db	2021	TTATGTTGGGACGAGAACCTGCTGCAATTGCTGCAATTGCTGCAATTGCTGCA	2080
Qy	2078	AGGACAAATGACTGTGAACTGCACTGGGAACTTGTGGGACCTTGTGGGACCTC	2137
Db	2091	AGGACAAATGACTGTGAACTGCACTGGGAACTTGTGGGACCTTGTGGGACCTC	2140
Qy	2138	AGAACGCTGAAATATGCTGAAATGTGGTGTAACTCTGGATAATTCTGCGA	2197
Db	2141	AGAACGCTGAAATATGCTGAAATGTGGTGTAACTCTGGATAATTCTGCGA	2200
Qy	2198	TGGCTTGGCTCCCTGACCCCTGCTGGCACTGGCTGCAACAACTGGGACCTACTCTGCTCA	2257
Db	2201	TGGCTTGGCTCCCTGACCCCTGCTGGCACTGGCTGCAACAACTGGGACCTACTCTGCTCA	2260
Qy	2258	TGGCTGCTCCCTGACCCCTGCTGGAGGGCCTGGCTGCAACAACTGGGACCTACTCTGCTCA	2317
Db	2261	TGGCTGCTCCCTGACCCCTGCTGGAGGGCCTGGCTGCAACAACTGGGACCTACTCTGCTCA	2320
Qy	2318	GGACTGTGAACCCGAGCTCAATTTCTGGACATTCTCAACACCAACTCAGCG	2377
Db	2321	GGACTGTGAACCCGAGCTCAATTTCTGGACATTCTCAACACCAACTCAGCG	2380
Qy	2378	TGCCCGAGGAATAACTGGACTGACTTGTGATGCTGAAACATACCCAGTGTAA	2497
Db	2441	TGCCCGAGGAATAACTGGACTGACTTGTGATGCTGAAACATACCCAGTGTAA	2500
Qy	2498	CAGANGATGTGAGGAGCTGGGGAGGTTCACTGGGACATTGTAACCTGGCAACTACCC	2557
Db	2501	CAGANGATGTGAGGAGCTGGGGAGGTTCACTGGGACATTGTAACCTGGCAACTACCC	2560
Qy	2558	AGGCCATTACCGGCCAACCCGAGTGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2617
Db	2561	AGGCCATTACCGGCCAACCCGAGTGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2620

2y	2618	CATCCCTGATGGTGGCCCCATGGAGATCTTCCCTGAGATCAGGGAGCACTGGGACTCT 2677	Db	3701	GAATTGAAATAATAAATTAAGAACCTTGAGACCA 1739
Db	2621	CATCCCTGATGGTGGCCCCATGGAGATCTTCCCTGAGATCAGGGAGCACTGGGACTCT 268C			
2y	2678	GTTGATGCGAAAACCTCTCATCAAATTCTGACAAATATAACCTGGAGACTA 273?		RESULT 5	
Db	2681	GTTGATGCGAAAACCTCTCATCAAATTCTGACAAATATAACCTGGAGACTA 274C		US-1C-198-846-11:92	
Qy	2738	CGAACGCCCATGCCTTCACTCCAGGTCAGAGCTGGATTCAAGAGCTGGACTCAA 2797		Sequence 1:192, Application US-/0198846	
Db	2741	CGAACGCCCATGCCTTCACTCCAGGTCAGAGCTGGATTCAAGAGCTGGACTCAA 280C		Publication No. US2003099974A1	
Qy	2798	TGAAGGGRAACAGCGCTAGAGGTTCAGAGTCCAGGTCACATATGAGGAGACTCA 2857		GENERAL INFORMATION:	
Db	2801	TGAAGGGRAACAGCGCTAGAGGTTCAGAGTCCAGGTCACATATGAGGAGACTCA 286C		APPLICANT: Lillie, James	
Qy	2858	GAACTCATGAGCATACTTCGAGTCAGGCTCATGATCTGAGAACCATCTGAGA 2917		APPLICANT: Xu, Yongyao	
Db	2861	GAACTCATGAGCATACTTCGAGTCAGGCTCATGATCTGAGAACCATCTGAGA 292C		APPLICANT: Wang, Yuzhen	
Qy	2918	AATACTTAAAGGATAGAAACTTATAAAGGCTCTGGTGTGATCTGGCCATCCCAGAA 2977		APPLICANT: Steimann, Kathleen	
Db	2921	AATACTTAAAGGATAGAAACTTATAAAGGCTCTGGTGTGATCTGGCCATCCCAGAA 2980		TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS	
Qy	2978	CTATTCAGTAGACAGCCAGGAGTCCCGAGATGTTTCAGATCTGTTACATCGATT 3037		TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND	
Db	2981	CTATTCAGTAGACAGCCAGGAGTCCCGAGATGTTTCAGATCTGTTACATCGATT 3040		FILE REFERENCE: NRI-C49	
Qy	3038	GCTACSTTCCAAGGTGTCAGGTTTGTGACCACTTACAGCCAGTGGCAC 3097		CURRENT APPLICATION NUMBER: US/1C/198-846	
Db	3041	GCTACSTTCCAAGGTGTCAGGTTTGTGACCACTTACAGCCAGTGGCAC 3100		CURRENT FILING DATE: 2002-07-18	
Qy	3098	TCAAATACAGATGGTGTGATAGGGTGGGAGAGGTGTCCTCTGGATGTCAGTCA 3157		PRIOR APPLICATION NUMBER: 6/0/306-220	
Db	3101	TCAAATACAGATGGTGTGATAGGGTGGGAGAGGTGTCCTCTGGATGTCAGTCA 3160		PRIOR FILING DATE: 2001-07-28	
Qy	3158	GCACAGTCGGTATATGCTGCTCCCTCTGGATGTCAGTCAATGGTGTGTCAGTCA 3217		NUMBER OF SEQ ID NOS: 14084	
Db	3161	GCACAGTCGGTATATGCTGCTCCCTCTGGATGTCAGTCAATGGTGTGTCAGTCA 3220		SOFTWARE: FastSEQ for Windows Version 4.0	
Qy	3218	TAATACAGATATTTCGTAATTGAACTTCTTCCAGCATGGTGTGATGATGACA 3277		SEQ ID NO: 11192	
Db	3222	TAATACAGATATTTCGTAATTGAACTTCTTCCAGCATGGTGTGATGATGACA 3280		LENGTH: 3091	
Qy	3278	CTGAGATGGTTGAGTCGATCTGCTGCTCCCTGTTGAGTCATGGTGTGTCAGTCA 3337		TYPE: DNA	
Db	3281	CTGAGATGGTTGAGTCGATCTGCTGCTCCCTGTTGAGTCATGGTGTGTCAGTCA 3340		FEATURE: misc_feature	
Qy	3338	TCACGGGCTGCCCTGAGCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 3397		NAME/KEY: misc_feature	
Db	3341	TCACGGGCTGCCCTGAGCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 3400		LOCATION: US/1C/198-846	
Qy	3398	GTCTTACTCCCTCAAGGACTCTGTAGTCAGGAAAGGGCCACAAATAGCTCTTAT 3457		LOCATION: 2978, 2979, 2980, 2981, 2982, 2983, 3085, 3086, 3087,	
Db	3401	GTCTTACTCCCTCAAGGACTCTGTAGTCAGGAAAGGGCCACAAATAGCTCTTAT 3460		LOCATION: 3086, 3089, 309C, 3C9,	
Qy	3458	CTGGAACCTCTGAGCT 3517		OTHER INFORMATION: n = A, T, C or G	
Db	3462	CTGGAACCTCTGAGCT 3520		US/1C/198-846--192	
Qy	3518	AGGCTCTGACAGGGCAACAGGGCAAGGGGAGAGGAGGAGGAGGAGGAGGAGG 3577		Query Match: 74.0%; Score: 2763; DB: 14; Length: 3091;	
Db	3521	AGGCTCTGACAGGGCAACAGGGCAAGGGGAGAGGAGGAGGAGGAGGAGGAGG 3580		Best Local Similarity: 96.3%; Pred: No. 0;	
Qy	3578	CACCCACCTTGAACCTGGAGGACTCTGCTTCCTCCAGCTCTCTCACTGCTG 3637		Mismatches: 0; Indels: 87; Gaps: 1;	
Db	3581	CACCCACCTTGAACCTGGAGGACTCTGCTTCCTCCAGCTCTCTCACTGCTG 364C		Matches: 263;	
Qy	3638	ACAAGGTTGATCCCAAGAAACTGAGTCAGGCTCTGAGCTGAGGAGGAGGAGG 3697		Conservative: 0;	
Db	3641	ACAAGGTTGATCCCAAGAAACTGAGTCAGGCTCTGAGCTGAGGAGGAGGAGG 370C		CGATGACTGTCAGGCTATATGCTGTTGAGCTGCTGTTGAGCTGCTGTTGAGCTG 375	
Qy	3698	GAATTGAAATAATAAAACTAAGGACTCTGGAGACA 3736		CGATGACTGTCAGGCTATATGCTGTTGAGCTGCTGTTGAGCTGCTGTTGAGCTG 376	

QY	917	GGGGCTGCTCATGGAAACCTGTGCTGTCACAAATGGAGGCTGTGACCTGTAAGGA	976	1910 CGACGTGGCTAAAAACCTCCAGAACATCTGAAAGCCAGGCGACTCTGTGGAGTGGC	969
Db	556	GCGGTGCTCATGGAAACCTGTGCTGTCACAAATGGAGGCTGTGACCTGTAAGGA	615	1636 CGACGTGGCTAAAAACCTCCAGAACATCTGAAAGCCAGGCGACTCTGTGGAGTGGC	685
QY	977	TACTCGACAGGTGTCACACTGCACTGTTCTGTGEGATTCACTTCAGTGGATGGAA	2336	1970 CGAGGTCTCATGGAAACCTCAATGGCTGCTGTCACAAATGGAGGCTGTGACCT	2029
Db	616	TACTCGACAGGTGTCACACTGCACTGTTCTGTGEGATTCACTTCAGTGGATGGAA	675	1696 CGAGGTCTCATGGAAACCTCAATGGCTGCTGTCACAAATGGAGGCTGTGACCT	1755
QY	1037	ZACATGTAANGATAATTGATGAGTGGCTGAGGTTGTGATCATTTCTGAA	1096	1973 CGAGAGACGCTGCATTATGCTGCTGTCACATGGCTGCTGTCACAAATGGAGG	2089
Db	676	ZACATGTAANGATAATTGATGAGTGGCTGAGGTTGTGATCATTTCTGAA	735	1756 CGAGAGACGCTGCATTATGCTGCTGTCACATGGCTGCTGTCACAAATGGAGG	1815
QY	1097	AAACATCGTGGCAGTTTGACTGTGAGTGGCTGACAGCTGCAAGAAGGTTAA	1156	2090 TTGTGAAACATGCCAAGACAGGAAATTCTGGGCCCCGAGCTGGAA	2149
Db	736	AAACATCGTGGCAGTTTGACTGTGAGTGGCTGACAGCTGCAAGAAGGTTAA	795	1816 TTGTGAAACATGCCAAGACAGGAAATTCTGGGCCCCGAGCTGGAA	1875
QY	1157	GAAGTCCTGGCAGAAGGTGGATGAGTGGCTGACAGCTGCAAGCAGCTGCA	1216	2150 TATGTCGAAATGTCGGGGCTGTGTCACACCTGTGAAATTCTGAGATGCTGAC	2239
Db	796	GAAGTCCTGGCAGAAGGTGGATGAGTGGCTGACAGCTGCAAGCAGCTGCA	955	1876 TATGTCGAAATGTCGGGGCTGTGTCACACCTGTGAAATTCTGAGATGCTGAC	1935
QY	1217	CAACACCCTGGCAGATTGGTGGATGAGTGGCTGACAGCTGCAAGCAGCTGCA	1276	2210 TTGCGCAGCTTGTGCTGCTGCGACCTTGTGCTGCTGAACTCTCTGTTCC	2239
Db	856	CAACACCCTGGCAGATTGGTGGATGAGTGGCTGACAGCTGCAAGCAGCTGCA	915	1936 TTGCGCAGCTTGTGCTGCTGCGACCTTGTGCTGCTGAACTCTCTGTTCC	1995
QY	1277	CACTCTGGAGACACCAATGAGTGGCATCAAAACAGGGCTGTGAGGCTGTGT	1336	2270 CTGTCGAGGAGGCTGTGCTGACAGCTGCAAGCAGCTGCAAGCAGCTGCA	2323
Db	916	CACTCTGGAGACACCAATGAGTGGCATCAAAACAGGGCTGTGAGGCTGTGT	975	1956 CTGTCGAGGAGGCTGTGCTGACAGCTGCAAGCAGCTGCAAGCAGCTGCA	2055
QY	1337	GAACACAGTGGCAGCTGATGAGTGGCTGACAGCTGGTACAGCTCACTGGAA	1396	2230 CAGAGTCACATGTCAGCTGGACATTCTTAACACACCAACTACCGATGTTAC	2389
Db	976	GAACACAGTGGCAGCTGATGAGTGGCTGACAGCTGGTACAGCTCACTGGAA	1035	2356 CAGAGTCACATGTCAGCTGGACATTCTTAACACACCAACTACCGATGTTAC	2115
QY	1397	AAAAGACTGTGGAAAGTGGTGGAGGCTTCCCTGCCACAAGSTGTCACCGCTG	1456	2390 CCAGTGSAGACATACGGCTGAAATAATTGTAAGGAAATTCTGCTCCAGGAA	2449
Db	1036	AAAAGACTGTGGAAAGTGGTGGAGGCTTCCCTGCCACAAGSTGTCACCGCTG	1095	2116 CCAGTGSAGACATACGGCTGAAATAATTGTAAGGAAATTCTGCTCCAGGAA	2175
QY	1457	GCACCTCGTAAAGACTGTGGAGGCTTCCCTGCCACAAGSTGTCACCGCTG	1516	2450 TACTAGACGACTCTGATGCTGCTCACAAACATACGGCTGAAATAAGGAGA	2509
Db	1096	GCACCTCGTAAAGACTGTGGAGGCTTCCCTGCCACAAGSTGTCACCGCTG	1155	2176 TACTACGACTGATGCTGCTGAACTTGTGCTGAAATAAGGAGA	2235
QY	1517	CCTCTCTTCA-----	1526	2510 AGGGGACTGSSAGATTTCTGGGTACATGGTAACTTGTGAACTTGTGAA	2569
Db	1156	CCTCTCTTCAAGACTGCAAGGCTCAAGGGCTACTCTGTCACCTCTCTCAG	1215	2236 AGGGGACTGSSAGATTTCTGGGTACATGGTAACTTGTGAACTTGTGAA	2295
QY	1527	-	1549	2510 AGCCAAACAACCGGCTGACCTTACCGCTGAACTTGTGAACTTGTGAA	2629
Db	1216	GAACAAACAAATAAATCAATGACTCTGCTTTRGGGATGTCACCATCAGACAG	1275	2296 AGCCAAACAACCGGCTGACCTTACCGCTGAACTTGTGAACTTGTGAA	2355
QY	1550	TGTAAACCTTAAAGCTAAATGAGGCAAGTGTGAAATGTCAGTGGCTTCGG	1609	2630 SGTCCTGAGATCTCTGGCCATAGGGAGACTGTGGGACTCTGGTGTGCGAA	2689
Db	1276	TGTAAACCTTAAAGCTAAATGAGGCAAGTGTGAAATGTCAGTGGCTTCGG	1335	2356 GTCCTCCGAGATCTCTGGCCATAGGGAGACTGTGGGACTCTGGTGTGCGAA	2415
QY	1610	GGTCTCTGGAGACAGCTACAGAGAGCAAGCTGCAAGTAAAGAGGTTCCCTAGT	1669	2690 AACCTCTTCACTCCAGTGTGCAACATAGAAACCTCTGGCCCTCAT	2749
Db	1336	GGTCTCTGGAGACAGCTACAGAGAGCAAGCTGCAAGTAAAGAGGTTCCCTAGT	1395	2416 AACCTCTTCACTCCAGTGTGCAACATAGAAACCTCTGGCCCTCAT	2475
QY	1730	TAAGGAAATGTTATCACTGTGAGTTGACTGAAACTAACCTAACCTAAC	1789	2810 CGTAGAGGCTTCCAGTGTGCAACATAGAAACCTCTGGCCCTCAT	2869
Db	1456	TAAGGAAATGTTATCACTGTGAGTTGACTGAAACTAACCTAACCTAAC	1515	2536 CGCTCTAACCTCAGTGTGCAACATAGAAACCTCTGGCCCTCAT	2695
QY	1790	TCTTGTGACTGCTGAGCTGCTGTCATGTCAGTGTGAGTTGACTGAAACTAAC	1649	2870 AGACATAGTTGAGTGGCTGCTATGCACTCTGAAACCATGAAATACTTAAGGA	2929
Db	1396	ANACCTTACATGCACTGCTGTCAGTGTGCAACATGTCAGTGTGCAAC	1455	2596 AGACATAGTTGAGTGGCTGCTATGCACTCTGAAACCATGAAATACTTAAGGA	2655
QY	1516	TCTTGTGACTGCTGAGCTGCTGTCATGTCAGTGTGCAACATGTCAGTGT	1575	2935 TAGAACCTTCAAGGCTGCTGAGCTGCTATGCACTCTGAAACCATGAAATACTTAAGGA	2993
QY	1850	CAGCTCTGAGAAAGGCCCTCACAGGGAGCAGTTCACTCCAGGCTCAGGCA	1909	2616 TAAGAATGCTACATGCTGAGCTGCTGTCATGCACTCTGAAACCATGAAATACTTAAGGA	2715
Db	1576	CAGCTCTGAGAAAGGCCCTCACAGGGAGCAGTTCACTCCAGGCTCAGGCA	1635	2933 CACGCCCAAGGAGCTCCGAGAGATGCTGAGCTGCTATGCACTCTGAAACCATGAAATACTTAAGGA	3049

```

Db 27-6 CACGCCAGGACTCCAGATGGTTCCAAGATCAGTGTCTCGAATGTCAGTCCAA 2775
Qy 3050 AGTGTCCGGTTTGGACCTTACAAATGACTGCGCACCGTGCACATACAAATG 3119
Db 2776 AGTGTCCGGTTTGGACCTTACAAATGACTGCGCACCGTGCACATACAAATG 2935
Qy 3110 TTCTGCTATAGGGTTGGGAGAGGTCTCTCTGCAATGCTGGGGT 3169
Db 2816 TTCTGCTATAGGGTTGGGGAGAGGTCTCTCTGCAATGCTGGGGT 2995
Qy 3172 ATTCGTGCTCCGGTATCAGTGACTCATTAGAGTTAAATTAAACGATAT 3229
Db 2896 ATTCGTGCTCCGGTATCAGTGACTCATTAGAGTTAAATTAAACGATAT 2955
Qy 3230 TTGGTAATTGA 3243
Db 2956 TTGGTAAAAAAA 2969

RESULT 6
US-09-898-570-11/c
; Sequence 11, Application US/09898570
; Patent No. US20020123612A1

; GENERAL INFORMATION:
; APPLICANT: GELACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA

; TITLE OF INVENTION: NOVEL HUMAN PROTEINS: POLYNUCLEOTIDES ENCODING THEM AND
; METHODS OF USING THE SAME
; FILE REFERENCE: 15366-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,.324
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
; LENGTH: 4821
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: PCIYX
; OTHER INFORMATION: 297832_B_0.707
US-09-898-570-11

Query Match 28.6%; Score 1069.2; DB 10; Length 4821;
Best Loca_Similarity 63.1%; Pred. No. 0;
Matches 1836; Conservative 0; Missmatches 993; _nDeleteS 81; Gaps 9;

```

Qy	196	CCGGGCGCCGAGGAAATTGTAGATGAGTCTGCGCAAGGGTAGACTGCCATGCCG	255
Db	4412	CCCTCTTCCCCTGCAGATGGATAGTGTGGAGGACTGAACTGCCATATCG	4353
Qy	256	ACGCCCTGTCTGAGAACACCCACCTCTTACAAGGTCCTCAGCCGCTCATCAAG	315
Db	4352	ATGGTATCGCCGAACACCCGAGGCTACATGAAAGTCATCGCAGTCAGTACAG	4293
Qy	316	GGAAAGSAGGCGGTGAGGACATCTATGAGTAAATTGAGCTCAATGAGGTGTTG	375
Db	4292	GGACGSSAAAACCTGCGAAAGCTGTTGATGAGTGGGATAATCGAGGTGTTG	4233
Qy	376	TCCATGACTGTGAATTTCAGGAAATTCTGTGTTGATGCTGTTGATGCTTGT	435
Db	4232	TGATGACTGTGTCACATCCCTGGCAATTACCGGTGATGTTGATGATTCACC	4173
Qy	436	TGGCTCATGAGGTCAATAATTGCTCTGAGCTGAGTCGAGAACAATGGGGCT	495
Db	4172	TGACATAAGGACACACTGTTGATGEGACAGTGAGTCGGGGCAACGGGGCT	4113
Qy	496	GCCAGCATACCTGTGTCACGTCATGGGAGCTATGAGTGTGCTGCAAGGGTTT	555
Db	4112	GTZAGCAGAGCTGTGTCACATGAGTGGGAGCTATGAGTGGGAGCTTCT	4053
Qy	556	TCTGTGAGTGCACATAGGACACCTGATTCAACGGCTGGTGAAGGGCTGAGCT	615
Db	4052	TCTCTAGGCAACACCAACATACCTGATTCCTGGCCAGACAGAGGAAATGATG	3993
Qy	616	ATAGGATCACGGCTTGTAGTCACATCTGCAAGGGGCCAACAGGGAGCTG	675
Db	3992	ACRAGAACACGGCTGCGCCACATTGCGGAGACCAAGGGGGTATGGCTGTG	3933
Qy	676	AGTCGAGGGCTGTTGAGCTGCCAGAACGAGACTGCACTTGTGAAACC	735
Db	3932	AATGCCGTCCTGCTCTGAGCTTACCAAACCAACGGGACTTAATTGACATGGAACT	3873
Qy	736	ATGGAAAGGTGGTGCAGGCTCTGTGAGCTAGCCATAGGGCATGGTGAAGCT	795
Db	3872	ATGGTAAAGGGCTGCGAGCACGGTGTGACAGACAGAGGCTGGGGTCCGGT	3813
Qy	796	GCCTTACAGTACAAAGTACAGATGGAGGACTGGCTTGTGAGGAGGAACTG	855
Db	3812	GCATATAAGGTTGCTTACACAGGGAAAGATGATCGGG-----3766	
Qy	856	TCTGAGGGTGAAGGAGAACACACATAGTGTGGATGCGGATAAACGGGTAAAC	915
Db	3765	---GAAAGGGCTTAGGACGACATCCCAC-----C-----3735	
Qy	916	GGGGCGCTGCTGTGAAAGTGTGCTCTGCTCAAAATGGAGGCTGTGCGAGG	975
Db	3734	AGGGCTTCTCATGAGCTGTGCTTCAAAACGGGCTGTGAGTAACTGCGAT	3675
Qy	976	ATACCTGACAGTGTGCAACTCGAGTGTCTGTGGATGCTTCACTCTCCAGT	1C35
Db	3674	ATGGCGACTGTGTCACAGCAGTGTGGCTCTGTGAGGCTGACATGAGGAA	3615
Qy	1036	AGCATGTAAGATAATTGATGAGTGGCTGAGCCGCAATTGAGGTTGTGATTC	1C95
Db	3614	AGAGTGTGAAAGATAATGAGTGGCTGCTTAACACAGCGGGGCTGTGACAT	3555
Qy	1096	AAACACATCTGGCAGTTGACTGCGCTGCAAGAAAGGTTAAATTATTAACAGAT	1155
Db	3554	GCACACAGTGGCAGCTGAGTGGCTGCAAGAAAGGTATAGCTCTCATGAT	3495
Qy	1156	AGAGTCTTGGCCAAAGATGGATGAGTGGCTCTGGATAGACCTGTGACACAG	1215
Db	3494	AGGGAAATGCGGAGATAAGGAGGTTGATGACCTGTTGACATGAGTGT	3435
Qy	1216	TCAACACCTGGCAACATTGCTGCTCTGCAACCGAGGTACACCCGTTGATGCT	1275
Db	3434	TCAACACACCGAAAGCTTCCAGTGTGCTCTGCTGATGTTGATGATATCA	3375
Qy	1276	CCSACTGTGGAGRCACCAATGAGTGGAGCACTGGCTCAGGGCTGTG	11335

3374 CCCACTGTGGGATGTGGATAATGAGCATCAACGGGGGGTGGCTGCA 3315
 1336 TGAACACAGTCGGCAGCTATGATGCGAGTCCACCTGG---GTACAAGTCRACTGAA 1392
 3314 TCAACATCTCTGGCACTTACAGTGTACCTGCCAACGGTGTGACTGGA 3255
 1393 ATAAAAGAAGACTGTGGAAAGTGAAGGGCTCTGGCCAAGTGTGTCCCCGGTGT 1452
 3254 ATGGCAAGATTGACAGGCCACTTAAGTCAGGGCACTCCCTGGCTCGAAAGCCA 3195
 1453 CCCTGACTGGTTGAGTGTAGGTGGGAGGGGTTCTAGATGAGTGACTGTGGAA 1512
 3194 TGCTCAAGTGTGGCAAGGCCACTTAAGTCAGGGCACTCCCTGGCTCGAGG 3125
 1513 TTACCCCTCTTCAGA-----TGTACCAACCATCAGGACAAAGTGTAACTTTAAGCTAA 1566
 3134 CGGATTPTTCAGGGTSCAGTGTGTCATTAAACACGGGCTCTTCAGATA 3075
 1567 ATGAGGCAAGTGTAGTTGAA-----ATATGCTTGGCTTCCGGGGTGTGCGAC 1620
 3074 AGGATCCAAATGCCSNTGCACTCCAGGAAACAAAGGCAACCTGGCAGAA 3015
 1621 CAGCACTACAGAGAACAGAACAGAACAGAACCTAACCTAACAT 1680
 3014 CCACAGGGCCAGGTGGCCCTGCTGTGATGCCAGGTACCCCTCATCACCTTAAGT 2955
 1681 GCAGETCTGGCAAGGCAAGTCCAGGGGCCAGGACCCCTAACGAAATGT 1740
 2954 G.GACTCTCTGGAAAGGGCAAGGGCGAGGGCGAGGGCAAGGGTCA 2895
 1741 TTA---TCATCTGTGAGTTGAGCTTGAACCAAAGGGTGAACGCTTGTGT 1797
 2894 CAAGGCTCACCTGGAACTGGGCAAGGTAGGGCAAGGCAACCCAGCTGCA 2835
 1798 ACCTGAGCTGATCTGGAAAGGAGGATGAAACTAACCAAAGGGTGAACGCTTGTGT 1857
 2834 GGCTGCCTGGCTCCGAGGAATGGAAACGGGGTGAAGGATCCCTGAAGATGCTCA 2775
 1858 GAAAGGGCTTCACAGGGGAGGTTCACCTCAAGCTCAGGGATGAACCTGAGCTGG 1917
 2774 GAAAGTCATTAACCGAACGCTGCTGGCTGAATGAGCTGAGCTGG 2715
 1918 CTAAAAGCTCCCGAAACATCTGAAAGCCATGAAAGGGTGGCAAGGCA-----GAGTCTGTGAGTGT 1968
 2714 CCCACAGGCCGGCTGGTAGGGAGGACAGGCAAGGCCATGGAGCTCTGTASGCCGG 2655
 1969 GCGAGGCTCAAGCAAACTGGTCAATTGTCAGTTGAGGGTGGAACTTATGATGGAG 2228
 2654 GGCGAACGGCTGGTGGACAAAGTGTCACTGCGCCGGAAAGCTTACACCGGCC 2595
 2029 CACGAGAACGCTGCAATTATGTCCAAATGAGGAAGGACAATGA 2086
 2594 AGACGAGAACGCTGTCGCACTGAGGAGGAACTTCAAGGAGGACGCT 2535
 2089 CCTGTGAACATGCCAACAGACCGGAAATTCTGGGCCCCCTGAAGGCCGAAAGCTGGAA 2148
 2534 CCTGCACCTTGTGCTGGGA-----GAGTCTGGCTCCAGGGGCTCTGGAGGCCAAC 2484
 2149 ATATGCTGATGCTGGAGGTGTGAGTGTGAACTGGTGTGCTGGCTGGAC 2208
 2483 ACGTCAACCTGGTCAAGGTAGTGTGCCAACCTGGCAACACTCTGTAGATGGTCAAGC 2424
 2209 CTGCGCAAGCTGTGCCCTGGCACGTTCAAGGCTGAGCTGTCGAACTTCCTTCCC 2268
 2423 CCTGTCAAGCTGCCAACGCTGAGGAGGAACTTCAAGGAGGACGCTATGCTCC 2364
 2269 CCTGTGGAGGAGGCCCTGCCAACAAACATCAGGAGGACTACTTCTTCAAGGAGGACTGTGAA 2328
 2363 CCTGTGTGGGCCCTCACCAAGCAAGTGAAGGGCCCATTCCTTCAGACTGACAA 2304
 2329 CGAGTTCAAATGTCACCTGGCACATTCAACACCAACACTAACGAGTATGTT 2388

Qy	1969	GCCAGGGCATGGAGAAACCAATGTCAGTGTGCAAGTGGCTGGGACCTATTATGATGAG	2028	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2654	GGCAGAACCGTGTGGACCAAGTGTGCAAGTGGCTGGCGAAGCTTACACGCC	2955	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2029	CACGAAACGGTGATTTATGTCAAATGGAACCTCCAAATGAGGAGGACAAATGA	2088	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2594	AGACGAGACTGTGCAAGGCCATGGCCATGGCCACCTTCCAGGAGGAGGAGCTT	2535	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2089	CTTGTAACCTGCCCAGGAAATTCTGGGCGCTTGAAACCCAGAAGCTTGA	2148	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2534	CCTGGCACCTTGCCCTGGGA-----GTGTGCCAACGGGCTTGTAGCCACA	2184	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2149	ATATGTCGAAATGTGGAGGTCGTGTCAAATCTGCAAGATGCTTGCAC	2208	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2483	ACGTACACCACCTGTGAGGTAGTGTGCCACTGGCAACACTCTGAGC	2124	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2209	CTTGCAGACTGTGTGCCCTGGCAAGCTGGCAACTCTCTGCTCC	2268	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2423	CCTGTAGCCTGCCACGTCACCTAACACCTGAAAGGGAGCACCTATGCTTC	2364	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2269	CCTGTAGGAGGCCCCCTGCCAACACATAGGGGCTACTTCCCTTCAGGACTGTGAA	2328	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2363	CTTGTGTGGGGCTTCAACCAACGGCATGAAGGGCCA-TTCCCTCGAGACTGTGACA	2304	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2329	CTAGAGTTCAATGTTACCTGGACATTCTACAAACCAACCTAACGATGTA-TCGT	2388	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2303	CCAAGTCCATGTGCTGCTCCAGGGCACTACTAACACCAACATCCACCTGTATTGCT	2244	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2389	GCCCCAGGGACACATAACCGCTGAAATTGTGTTCTGCCAGGA	2448	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2243	CTGCCTATGGGCTCATACGCCGACTTCCTCAGACCTTCGAGCTTCAGGAA	2184	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2449	ATACTAGACTGACTTGTGCTCCACAAATACCCAGTAAACAGAGATG	2508	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2183	ACACAAAGCACAGACATGTGATGCTTACCAAGTGTGCCCAATGCAAGATGTCAGTGTG	2124	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2509	QAGGGAGCTGGAGATTCTACTGGGATTCATGGAATCCAAACTAACCTGAAATTAC	2558	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2123	CTGGGAGCTGGTGTGTTCTGGCTATATGAGTCCTGCTCATGAGTACTACC	2064	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2569	CAGCCAAACACCCGATCACTGGGACCATCAACCCACCCCCAAGGGCGAACCTGTGATGCG	2628	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2063	CAGCTGTGTGGATGTCATCTGGAAACATCAACCCCCAACGGCAGATCCATG	2034	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2629	TGGTCCCTGAGAACTCTCTGCCATAGGGGACTATCTGGTGGGACTATCTGGTGTGCGGA	2638	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2003	TGGTACCAAGATCTCTGGCTCATCTGGCATCTGGATGAGTGTGGGACCTCTGTCATGAGA	1944	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2689	AAACCTCTCTCAATCTCTGACACATGAACTTCTGACACATGAACTCTCCAGAACCTGAGCCCA	2748	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	1943	AGAACATCCATCCCCATCTCTCAATCTGAAACATCAACCCCCAACGACCTGAGCCATG	1884	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2749	TGGCTCTACCTCACTGGTCAAGTGGTGAATGTCATGAAAGGAAACA	2588	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	1883	TGGCTCTACGCCCTCTCTCACTGGTCAAGTGGTGAATGTCATGAAAGGAAACA	1824	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2809	GGGCTAGAGGGTCCAGGCTCATAGTGACATATGAGGACTACAGGAACCTATG	2858	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	1823	GGGCCGTTGGCTCCAGATTCCTCATGTCATGAGCTATGAGCTTGTGTTGATG	2764	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2869	AGACATAGTGTGGAGATGGGGCTCATGATCTGAGAACATCAGGAAATTAAAG	2928	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	1763	AGACATTGTGGAGATGGGGCTCATGCTCTGAAACCCAGGAGATTAAAG	1744	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2929	ATAGAAACATTACGGGCTGTGCTGATGTCATGCTGCTATGTCATGAGTACTTCAGT	2988	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	1703	ACAAGAACGTCATCAAGGCCCTCTGGTGTGAGGCTAGCCACCCAGAACATTCTCAAGT	1644	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557
Qy	2989	ACACAGGCCAGGAGTCCCGAGAGATSTTCCAGATGCTGATGCTAGTCA	2048	Qy	3049	AAGTGTCAAGGTTTGAACCTTACAAT	3078
Db	2643	ACACAG---AGAACACAGGAGATGCTGCCAAAATCTCAAGCTCCTGCTCA	1887	Db	1586	AAGTTCAGCTTCAGGCCCTAACAT	1557

1 LENGTH: 4556 ; TYPE: DNA
 2 ORGANISM: Unknown Organism
 3 FEATURE:
 4 OTHER INFORMATION: Description of Unknown Organism: PCIYX
 5 OTHER INFORMATION: 297832_B_2
 6 US-09-839-446-13

7 Query Match 26.6% ; Score: 1067.6; DE 11; Length: 4550;
 8 Best Loca. Similarity 63.1%; Pred. No. 3; Mismatches 999; Index 82; Gaps 10;
 9 Matches 1849; Conservative 0; Description of Unknown Organism: PCIYX
 10 OTHER INFORMATION: 297832_B_2

11 2468 CAACTACCAGGCAACTACCCAGCTGGAGTCATCTGGACATCAACCCCCAAC 2527
 12 2609 CAAGCCGCATCCTCATGTCGTCCTGAGATCCTGCCATAGAGGAGAATGTGCG 2668
 13 2528 CAAGCGAAGATCCTATGGTGTACAGAGATCTCCATGAGATGATGTGCG 2567
 14 2669 GGACTATCTGCTGATGGAAAACCTCTTCATCCATTCTGACAACATATGAAACCTG 2728
 15 2588 GGACGTCCTCGTCATGAAAGAACATCCTCCATTAACCATATGAGACCTG 2647
 16 2729 CGAGCTTACAGACGCCCATGGCTTACCTCCAGTGAAAGAGCTGATGTGAGT 2788
 17 2648 CCAGACCTTACAGGCTCCATGCTGCTTACGGAGCTCGTGGAG 2707
 18 2789 CAAGTCATAGAGGAACAGCGCTTACGGGTTCCATACGTGACATATGATA 2848
 19 2708 CAAGAAGGGCAACAGGCTTACGGGAAAGGGAGGGAGGGAGGGAGGGAGG 2767
 20 2849 GGACTACCGGAAACTCATGGAGACATAGTTCGAGATGGCAGGCTTATGCACTGAGGA 2968
 21 2768 GGACTATGAGGAGCTGGAGACATGGTGGAGATGGGGCTTATGCTGAA 2827
 22 2909 CCATCAGGAAACTTAAAGGATAGAACCTATGAGGCTTACGGCTCCGCA 2968
 23 2828 CCACCGGGAGATTAAAGGCAAGAAGCTTACGGGCTTACGGGCTTACGGCTTAC 2897
 24 2969 TCCCAGAAGACTTAAAGGATAGAACCTATGAGGCTTACGGCTCCGCA 3029
 25 2888 CCCCGTAACTTAAAGGATAGAACCTATGAGGCTTACGGCTCCGCA 2944
 26 3029 CATCGATTTGCTCAAGTTCAGGTTTGGAGCTTACAAAT 3078
 27 2945 CATCAGCTGTCCTCGTCATCAAGTTCCAGETTCTGGCCCTACAAAT 2994

28 RESULT 9
 29 1 Sequence 13, Application US/09839446
 30 ; Publication No. US0030050232A1
 31 GENERAL INFORMATION:
 32 ; APPLICANT: GFRLACH, VALERIE L.
 33 ; APPLICANT: ELLERMAN, KAREN
 34 ; MACDOUGALL, JOHN R.
 35 ; SMITHSON, GLENDA
 36 TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
 37 TITLE OF INVENTION: METHODS OF USING THE SAME
 38 FILE REFERENCE: 159.66-776
 39 CURRENT APPLICATION NUMBER: US/09/839,446
 40 CURRENT FILING DATE: 2000-04-19
 41 PRIOR APPLICATION NUMBER: 60/198,293
 42 PRIOR FILING DATE: 2000-04-19
 43 PRIOR APPLICATION NUMBER: 60/198,645
 44 PRIOR FILING DATE: 2000-04-20
 45 PRIOR APPLICATION NUMBER: 60/210,809
 46 PRIOR FILING DATE: 2000-06-09
 47 PRIOR APPLICATION NUMBER: 60/199,476
 48 PRIOR FILING DATE: 2000-04-26
 49 PRIOR APPLICATION NUMBER: 60/200,025
 50 PRIOR FILING DATE: 2000-04-26
 51 PRIOR APPLICATION NUMBER: 60/224,610
 52 PRIOR FILING DATE: 2000-08-11
 53 PRIOR APPLICATION NUMBER: 60/200,024
 54 PRIOR FILING DATE: 2000-04-26
 55 PRIOR APPLICATION NUMBER: 60/199,880
 56 PRIOR FILING DATE: 2000-04-26
 57 PRIOR APPLICATION NUMBER: 60/218,591
 58 PRIOR FILING DATE: 2000-07-17
 59 PRIOR APPLICATION NUMBER: 60/271,814
 60 NUMBER OF SEQ ID NOS: 49
 61 SOFTWARE: PatentIn Ver. 2.1
 62 SEQ ID NO 13
 63 1076 AGTTGATCATCA-TTCCTGAAACATGTTGAGGCTTGGAGGAGGCTTGGAGGAGG 1135
 64 1016 CACTTCCAGTGGATGGAGGAGCTTAAGGATTTGATGCTGCGGAGCCCTAAATGGS 1075
 65 916 -
 66 956 CCTGACCCGACCTGTTGGATACCTGGCTTACAGTGGTCAACTGGATCTCTGGATT 1015
 67 957 CTGAGACAGTGGCTGATGAGCTGAGCTGACTGGATCTGGTCAACTGGTCTGGTT 916

RESULT_11

Qy	2984	CAAGTACAGGCCAGGGTCCGAGGATGTTCCAAGATGTTAACAAATGAC	3043
Db	2922	CAAGTACAGGCCAGGGTCCAGGAATCCAGGATGTTCCAGGTCTTACAATGAC	2981
Qy	3044	TTCGAAAGTGTCCAGGTTTGAGACCTTACAATGAC	3081
Db	2982	TTCGAAAGTGTCCAGGTTTGAGACCTTACAATGAC	3019

Sequence 1; Application US/09855824
 Publication No. US201030166048A1
 GENERAL INFORMATION:
 APPLICANT: YAN CHUNHUA et al.
 TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
 NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
 TITLE OF INVENTION: USES THEREOF
 FILE REFERENCE: CIC01233
 CURRENT APPLICATION NUMBER: US/09/09855,824
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NC_1
 LENGTH: 3877
 TYPE: DNA
 ORGANISM: Human
 US-39-09855-624_1

Query Match 28.2%; Score 1054.8; DB 12; Length 3877;
 Best Local Similarity 62.3%; Pred No. 9.5e-312; Mismatches 962; Indels 207; Gaps 9;
 Matches 1923; Conservative 1923;

Qy	89	CSCGGGCCAAACGTCGGGGCTGGCGGCTGCTGCTGCTGCTGCTGCTGC	148
Db	48	CSCAGCGCCAGGGCCGGCATTGAGATGGCGAGCGAGCGAGCGAGCG	107
Qy	149	GCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC	208
Db	108	CCTGTTGCTGGCTGGCAACGGGGCGGGTGGCGGGGGCAGGGCTCCAGG	167
Qy	209	GGAGGATGAGATGAGTGTCCAAAGGGCTAGATGACTGCTGCGACGCC	268
Db	162	AGTCGAGTGTGGAGTGTAGGGGACAAATGATGCACTATGCACTGTG	227
Qy	263	GAACACCCACCCCTTACAATGCTCTGGCAAGCTACACGGGAAAGCGAG	328
Db	229	GAACACGGCCAAAGCTTAAATGCCCTGAGGCGTACACGGGAAAGCGAG	287
Qy	329	GTGTAGGGCATGATGAATGTTGAAATG--AGCTTAATGGGGTGTGCTG	365
Db	269	GTGTAGGGCATGATGAGGTGAGAATGACTACTAAATGGGGTGTGCTCAG	347
Qy	386	TTGATATTCCGGAAATTGCTGTCATGTTGAGCTTGTGCTTGTGCTCA	445
Db	348	CATCACATCCGGGAACTACAGTGTTGCTGCTTGTGCTTGTGCTCA	407
Qy	446	CGCTGATAATTGTCATGTTGAGCTGAGCTGCTGCAACATGGGGCTGCA	505
Db	408	TGGAAACATGCTGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAG	467
Qy	506	CTGTGTCAGCTGATGGGGAGGTATGGGGTGTGCTGCTGCTGCTGCTG	565
Db	468	CTGCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	527
Qy	566	CAAATGAGAACCTGCAATCCGGCTGAAAGGGCTGACTGTGATAAGGAT	625
Db	528	AAAGCAGCATACCTGATCCACCGCTCAATAGGGTATGAACTGCAAGAC	587
Qy	626	CGCTGTAGTCATCTGCAAGGAGGCCAACGGGAGCTGCTGCTGCTG	685
Db	588	TGCTGTGCCACCTGCTGGAGAATGGCGAAAGGCTGCTGCTGCTGCTG	647

QY	686	TGGTTTGAGCTGGCAAGAACCGAGGAACTGGAGACTGCATCTGTAACCATGGGAACGG	745	1567 CGTAAACCTTACATGAGGCTCTGCAAGCAAGTCCAGGAGCCCTGGCACCAAGGAC	1726
Db	648	GGCTTGTACCTGGCCAAMCAGAGGACTGACACTAACCTGAAACG	707	1638 GCTTTCGACCTGAGTGTGACTCTCAGTGTGAACTCCAGAGGGCCCTGGCAAGTCCC	1697
QY	746	TGGGTGCCAGGACTCTGTGAGGATAACAGCCGATGGCCAGAGTGCCTCCACA	905	1727 CCTTAAGGAAAATGTTTATGACTGTGTTGAGTTGAGCTTGAACAAACTAACCAAAAGGAGT	1783
Db	708	GGGTGCCAAGACAGTGTGGGACAGACAGCCCCAACGGTGGTCCAGAGA	767	1638 ATCCAGGGGTGTCGCACTACAGAGGAGTTGAGATCGAGAACAGGAGGC	1757
QY	906	GTCACAGATGGACACAGATGGAGGCTGCTTSGCAGAGGACACTGCTGGAGGT	965	1784 GACGCTTCTTGTGACTTGAAGCTGATGCTGAGCCAAACGAGAGGGCTCGTAAAGG	1843
Db	768	GTACGCCCTCAACTCGAGGTGACGTG	797	1758 CTCAGACACATCGAAAGGGACTGCTGGAGGAGGAGAACAGGCTCGAGGGC	1817
QY	866	GACAGAGGAAACACACATGAGTGGTGGATGGGATAAAACGGGTCAAAGGCTGCTGCT	925	1844 CATCGCAGCGCTGAGAAAGGCTGACAGGAGGAGTTCACCTCAGCTCAGGCT	1903
Db	798	-	797	1818 CATCAAGKCCCTCGCGDAAGTGGCATGGCCAGCAGTCTATGTCAGGCTCTGGCAC	1877
QY	926	CATGGAAACGTGCTGTCACAATGGGGCTGTGACCTGACCTGTAAGGATAACTTCGAC	985	1934 GAACCTCACGTGAGCTGCTAAACACCTCCAGAGACATCTGAAAGCAGGAGTCCTGGG	1963
Db	793	CATCGAGACGCTGCGGAGTCATAAACGGGAGACATCRAAGGACACAGGCC	857	1818 TGAGTACAGGTAGCCAGGAGGAGGAGGGGG--CATGTTGGGGAGGGGG	1934
QY	986	AGTGGTCACAGCAGTGTCTGTTGAGTACTCTCAGTGGATGGAAAGAGCTGTA	1145	1934 AGTGGCAGGGGTCTCATCGAGGAAACAAATGAGCTTGTCAATTGAGCTTGGAC	2023
Db	858	TGGCGTGGATGCGAGTCAGTCGCCCCGTTGATGAGTCACCTGCGGAGACATCRAAGG	917	1935 CGAGGGCAGGTGCTCATCGAGGAGACATGAGCTTGTGCTGCGGACCCACTTCG	1994
QY	1046	AGATATTGATGAGTGGCAGACGCCGAAATGGGGTTGTGATAATTCTGCAAAACATCTG	1105	2024 TGAGGAGGAGGAGCTGCTCATTTGTCCAAATGSAACCTTCGAAATGAGGAGCA	2083
Db	918	AGACATCAACCAGTGGCTGAAACAGGGGCTGCGACACTCTGCGCACACCC	977	1995 TGGTAGGCTCGGSCAGCTGCTGCTCATATGAGGAACTACCCGAGGAGCA	2054
QY	1106	GGGAGTTTGACTGCGCTGAGAAGGTTAAATTATAACAGATGAGATGCTG	1155	2084 AATGACTTGTGACCATGCGGAGGAGGAAATTCCTGGGCTCTGAAAGCCCCAGAGC	2443
Db	978	GGGAGCTTCAAGTGGCTGGCGGGAGGGTACAGCTGGTCAAGGAGGACCTG	1037	2055 GCTCAGTTGSCACACCGTGGCTGGCTGGCTGCTGCTGCTG	2105
QY	1166	CCAAGATGCTGATGAGCTGCTTGGATAGGACCTGTGACCAACCTGCATCACCC	1225	2114 TTGAAATTGTGCTGAGTGGCTGAGCTGAACTCTGGAATATCTGAGATGCTT	2203
Db	1038	CCAGGACATCCAGACTGCTCTTCAGGCGAACCTGTGACCATCTGCAACTCTCCC	1097	2105 CCSCRAAGCTGCTGAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	2165
QY	1226	TGGCACATTGTTGCTGCTGCTGAAACCGAGGTACACCCCTGATGCTCACCC	1285	2204 TGACCTTGCCAGCTCTGCTGCTGAGCTGAACTCTGAGCTGAACTCTG	2263
Db	1098	GGGAGCTTCAAGTGGCTGCTGCTGCTGAGGACAAACCCACTGCG	1157	2166 CAAGCCCTGGCCAGGCTGCCCCTGGCTGAGGAGGCTGAGGAGGAGGAGG	2225
QY	1286	AGACACAAATTAGTGTGAGCATCAACACGGGGCTGCAAGGGCTGCTGTAACACAGT	1345	2264 CTTCCTCTGTGGAGGAGGCTGCTGAGGAGGCTGCTGAGACTCTTCAGACTG	2323
Db	1158	AGATGTGGACAGTGGAGCTGAGCTGAGAACGGCTGCTCACACCA	1217	2226 CTTCCTCTGTGGAGGAGGTTGCTCACCAACAGGAGGAGGAGGAGGAGG	2285
QY	1346	GGGAGCTTATGATGAGCTGCTTGGATAGGACCTGTGACCAACCTGCATCACCA	1455	2324 TGAACCCAGGAGTCAATTGCTGAGGAGGAGGAGGAGGAGGAGGAGG	2393
Db	1218	GGCAGACTACAGTGGCTGCTGCTCCCGGGCTAACCTGAAACGGGAAGGATTG	1277	2286 CGGGCTAAAGTCAGCTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCT	2345
QY	1466	TGTGGAA--AGTGAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1462	2384 TGTGTCGCGCAGCTGCGAAACATAACGGCTGAAATTGGAAATAATTTGTTCTGCC	2443
Db	1278	CCTGGAGACGCCAAGTGTCTTCTGGCCAGACCTCCGGCCAGCTGTCCTG	1337	2346 CCCTCCAGCAACCCSAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2405
QY	1463	CCTGTAAGAGTGTGGAGGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1522	2444 AGAAAATTTACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG	2503
Db	1338	ZAGCAGCGAGGCGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1397	2406 GGCGAACACGAGAGCTGGCTCCACCAACGAGCTGGCTGCTGCTGCTGCTG	2465
QY	1523	ZTCAGATG-----	1510	2504 ATGGAGGGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2563
Db	1398	GCAGACTGGAAAAATAGCTAGTGTGCTGAGGTTCAGGGCCGGAGGGAAAGGC	1457	2466 CTGGGGCGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2525
QY	1531	-----	-----	2564 TACCCAGCAACCCSAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2623
Db	1458	GCTGCGAGAACGCCAGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1558	2565 TACCCAGCAACCCSAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2585
QY	1559	TAGCTTAATGAGGAGGAGTGTGTTGAA-----AAATGCTGAGCTGTTGAGG	1517	2624 GATCGTGTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2683
Db	1618	ZACCCCCATCACACAGAGGGCCGCTGCTGAGCTGCTGCTGCTGCTGCTG	1622	2586 CATGCTGTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2645
QY	1613	TCTGCGACCAGGACTACAGGAGAACGACAGCT-----CAGTAAGAGGAGCTGCTG	1666	2684 GCCTAAACCTCTTCATCAAATGAACTGACACATGAACTGAACTGAACTGAA	2743
Db	1578	CCACAGCCAGGACGAGCAAAAGGAGCAGGCCAGGAGACCTGCTGCTGCTG	1637	2616 GAGGAAGNTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2705
QY	2744	CCCCATGGCTCACCTCACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2803		

RESULT 12
US-10-195-142-3
; Sequence 3, Application US/10195142
; Publication No. US2003036163A1:
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; ATTORNEY: Wettstein, Daniel Albert
; ATTORNEY: Mauck, Kimberly A
; TITLE OF INVENTION: NOVEL PN9826 NUCLEIC ACIDS AND USE THEREOF
; FILE REFERENCE: 1051_01
; CURRENT APPLICATION NUMBER: US/10/195_142
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,323
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-195-142-3

Query Match Score 1052_2; DS 14; Length 2967;
Best Local Similarity 62.7%; P-reqd. No. 5e-311; Mismatches 923; Indels 207; Gaps 9;
Matches 1903; Conservative 0; Pairs 10;

Qy	153	CTGCTGTGCGGGGGCCCTCCGGGCGCTCCGGGCGCTCGGCGACAGCTCTGGATCAATTCGAAGG 2765
Db	2706	GCCATGGCTCACCTCCCCCTGGGAACTGGTCCAGTCAGTCAGTCAATTCGAAGG 2765
Qy	2804	GAACAGGGCTGAGGGTTCCGGTCCCATAGTGACATATGAGGACTACCCAGGAACT 2863
Db	2766	GAACAGGGCAAGGTTCCAGTGCCCTATGTCACCTCAATGAGGACTACCCAGGAACT 2855
Qy	2864	TATTGAGAACATAGTTCGAGATGGCGGGCTATGATCTGAGAACATCAGGAAATAAT 2923
Db	2826	CATAGGGACATCGTGGCGATGGGCCCTTAAGGCACCAAGGAAATT 2885
Qy	2924	TAAGGATAAGAACATATCAGGGCTGTTGATGTCCTGGCCATCCCAGAACATCT 2933
Db	2886	GAAGAACAGGAGCTGATCAGGCCCTCTCTGACGTGGCTGGCATCCAGAACATCT 2945
Qy	2984	CAAGTACAGGCCAGGGATGTCCTGGAGAGATTTCCAAAGATGTTCATCGATTCTAC 3043
Db	2946	CAAGTACAGGCCAGAAATCAAGGATGTTCCACGCTTCATCAAATGCTGEG 3005
Qy	3044	TTCCAAGTGTCCAGGTTTGGACCTTACAAATGAC 3081
Db	3006	CTCCAAAGTGTCTGGCTCCAGGCTTACAAATGAC 3043

513 CTAAAGTCATGGGAGCTATGAGTGTGCAAGGGGGTTTTCCTGAGTGACAT 569
2c 397 GTAAATGSCATGGCACCTTCAAGTGTGAGCTTCAGTGGCTTCCTTAGTGACAC 456
Qy 570 CAGCACACTGATTCAGCTGGAGGGCTGAGCTGATAGGATAGCTGAGTCAGGC 629
2c 457 CAGCATACCTGATCCAGCGCCAAATGAGGGTATGAACTGATGAGAACATGGC 516
Dy 630 TTAGTCACATCTGCAAGGGGCCAGGGCTGGCTTGTAGTGCAGGCCCTGT 689
Qy 517 TGGCCCATCTGCGGAGGCGCCAAAGTGGCTGGCTGAGCTGAGGCCGGC 576
Dy 690 TTAGGCCGGCCAGAACAGGAGAACAGAACATCTGACCTGACCTGAACTGG 749
Qy 577 TTGACCTGCCCAGAACAGGAGGATGCAACTAACCTAACCTGAACTGGGGC 636
Dy 750 TGCGAGCTCTGTGAGGATCAGCCATGGCCAGGGTGAAGCTGAGCTGACATGAC 809
Qy 637 TGCGAGCTGAGCTGGAGAACAGAACAGGAGAACAGGCCCGGAGGCTGGTGTGAC 696
Dy 810 AAGATGCAACACAGATGGAGGAGCTGCTCTGGGAGAGAACATGTCCTGGAGGTGACA 869
Dy 697 GCCTCCACTCTACAGGTCAGCTGAC 722
Qy 870 GAGGCAACACACATGTCAGTGGATTAACGGGTAACGGGTTGGCTGATG 929
Dy 723 -----
Dy 930 GAACGTGTGCTGCAATGGAGGCTGTGACCGACACTGTGAGGATACTCGACAGT 989
Dy 727 GAGACGTGCGCACTCAAAACGGGCACTGCGACGGCACTGCACTGCG 786
Qy 99C GTGCCAGCTGGAGTGTCCGTGTTGGATCTCAGTGGATGGAGAACATGTAAGAT 1049
Dy 787 GTGGGATGAGCTGGCTCCGGTTGGATTTCTACAGCAGGGAGGAGAACATGCAARAG 846
Qy 1050 ATGGATGACTGCCAGACCCTGAGGGCTGATCATTTCTGCAAAAACATCGTGGC 109
Dy 847 ATTCAGGACTGCTGGTAACAGGGGCTGGGACACTCTGGGAACTCTGGGAC 906
Qy 1110 AGTTTGACTGCGCTGAGGAAAGGTTAAATAATACAGATGAGAATCTTGGCAA 1169
Dy 307 ACCTTGAGTGCCTGCGGCTGCGGAGGGCTACACAGCTGCTCACCGAC 906
Qy 1170 GATGGGAGTGGCTCTGGATGGACCTGACCTGCTCATACCCCTGGCG 1229
Dy 367 GACATGAGGACTCTCTGAGGACCTGCTGAGGACCTGCTCACATGCTGAC 366
Qy 1230 ACATTTGCTGTGTTGCAACGGGCTACACCTGCTGCTGAGGAC 1289
Dy 1227 AGTCACATGCTGTCACCGGGTACACTGCTGGGAT 1056
Qy 1290 ACCATGAGTGGACCATAACACGGGAGGCTGAGGCTGAGCTGCTGGCG 1349
Dy 1287 GTGAGACGACTGACCATSAGAACGGGAGCTGACCTGGGAAACACAAAGGC 1146
Qy 1350 AGCTATGATGCGCTGGTACCTGGTACAGAAGCTGAGGCTGAGGCTGAGG 1409
Dy 1147 AGCTACGACTGCTGCTGCTGGGGGGCTCACCTGGAAAGGATGGGTG 1206
Qy 1410 GA---AGTGAAGGGGCTCTGCCAACAGTGTGCTGACCTGGCT 1466
Dy 1227 GAGACGGAAAGTCTCTGCTGAGCTCCGGCCAGCTGGTCTGAGTC 1266
Qy 1467 AAGGTGTGGAGGAGGGCTCTCAGTGTGACTCTGGCATCTCTCTCA 1526
Dy 1267 AAGCAGGGGTGAGCTCTCTCTGCGGGCTCACACCTCTCGGGCA 1326
Qy 1527 GA-----
Dy 1327 GACTGGAAATAATGCTAACTGCTGAGCTGAGGCTGAGGCGCTG 1386
Qy 450 CATAATGTCATGGTATGGGAGGTGCTGGAGAACATGGGGCTGCCCCATACCTGT 509
Dy 337 CACAACTCCCTGGATGAGCTAACATGGGAACTACGGTCAAGGAAAGGTG 396

Qy	1529	- - - - -	TGTCAACCATAGAACAGTGTAACTTGGGAGGGTCTGT	AAGTAAAGCTTGATGTTCCGAGGGTCTGT	TAG 1567
Db	1387	CAGAAAGCAAAGGCCACAGCTGGCTCGGCTCAGTCACACC	1446		
Qy	1563	C ^{TAAT} GTAGGGAAAGTGTAGTTGAAC	- - - - -	AAATGCTGAGCTGTTCGGAGGGTCTGT	1616
Db	1447	CCATCAAACAGGGCCGCTCAATCCAGATGCCAATGCCCCAC	1506		
Qy	1617	GACCCAGCACTACCAAGAGAACAGT	- - - - -	CACTAAAGAGGACTCCGTCAGTA	1673
Db	1507	AGCCAGGCACGACCAAGGACCCAGGACGGCAGCGAC	1566		
Qy	1671	AACTTATCATGAGCTCTGAGAAGTCCAGGACAACTTGGC	1730		
Db	1567	TTCGTCACCCCTCAAGTGTGACTCTCTGAGAAGGGCC	1626		
Qy	1731	AAGGAATGT - - TTAATGACTGTGAGTTGAGCTGAACTAAC	AAAAGGAGGTGACA	1787	
Db	1627	AAGGGGTTCCACATACAGAGTGTGAGATCAGACAAGG	AAAGGGCTCA	1685	
Qy	1788	GCTCTCTGTGACTGTGAGCTGATCTAAAGGGAAACCGAGA	ANGCGGCTCCGTAAGGCCAT	1847	
Db	1687	GACACATGGCAACCGGACTGTGTTGGAGGACAGCAG	CTGAGGCCCATC	1746	
Qy	1848	CGACGCTCTGAGAAGGGCTCACAGGGAGCAGTTACCT	CCAGGCTCAGGGATGAAAC	1907	
Db	1747	AAGACCTGGCAAGTCTATGGCCGAGCTTCTATGTCAGG	CTGAGCTGAG 1806		
Qy	1908	CTGCACCTGGCTAAAAAGCCTCCCAGAACATCTGAAC	CCAGGCAAGTCTGTGGAGT	1967	
Db	1807	TAGAGGTAGGCCAGGGCAGCAAGGGCTGAGGGCAG	--GGGGCTGTGGGCA	1863	
Qy	1968	GGCAGGGTCATGCAAGAACCAATGTCAGTGGCTGGAC	CTATTATGAGGAA	2027	
Db	1864	GGCAGGGTGTACAGGACAAATGCTGTCCTGCGCC	TGGCAACCCACTCTGGTGT	1923	
Qy	2028	GCGAGAACGCTGCAATTATGTCATAATGAACTCTCC	AAAATGGGAAGGACAAATG	2087	
Db	1924	GAATCTGGCCAGTGTGTCATGTCAGTGGCTGGAC	ATGGAGGGGCTC	1983	
Qy	2088	ACTTGTGAACTGTGCCCTGAACTTCTGGGCTCTGAAG	ACCCAGAAAGCTGG	2147	
Db	1984	AGTTGCAACCTGGCCAGCACGCCAGGGCTGGCTG	-- -- -- --	CCGGTGGCGC 2034	
Qy	2148	AATATGCTGTAATGCTGGGACTTCTGAACTCTGTGAA	TATCTGCAAGATGGCTTGCTG	2207	
Db	2035	AACGGTGTGGAAATGGGGCTGGCTCTCTGGCTG	GGATGGCTTGCA	2094	
Qy	2208	CCRTGGCCAGCTGGCACTACAGGCTGAACTTGGAA	ACTCTGGCTTC	2267	
Db	2095	CCCTGGCCAGGGCTGGCACTACAGGCTGGCA	CTGGCAACCCGGCTG	2154	
Qy	2268	CCCTGTGGAGGAGCTCTGGCAACAACTACAGGAGT	ACTCTGGCTGAA	2327	
Db	2155	CCCTGTGGAGGGCTTGGCTACAAACACCAACTAC	CCGGTGTGGAG 2214		
Qy	2328	ACCAAGTTCAAATGTCACCTGGACATTCTGAA	ACCCAGCTTCTGGACTTG	2387	
Db	2215	GCTAAAGTGCACGTGCTCCCGGCAACCACTAC	ACACCAACCCGGCTGCA	2274	
Qy	2388	TGCCCACTGGAACTACAGGCTGAAATTGGAAA	AAATAATGTTGTTCTGGCCAGGA	2447	
Db	2275	TGCCCGTGGCACTACAGGCCAGTGTGGCA	ACCTGTGCTGGGC	2334	
Qy	2448	AATACTACGACTGACTTTGATGGTCCA	AAAATAAACAGAAATG	2507	
Db	2335	AAC2CAGGACAGACTTCAGGCTCAGAACGAA	AAACAGGCTGCA	2394	
Qy	2508	GGAGGGGACTGGGAGTTCACTGGTACATGGTAA	TCCCAAGGGCAATTAC	2567	
Db	2395	GGCCGGCAGGCTGTGACTACACGGTACATGGT	ACCTGGCCCAACTACCTGGCACTAC	2454	
Qy	2568	CCAGGCAACACCGAGSTGAGTGGACCA	CTAAACCAACCCCCAAAGCGEGCA	2627	

Db	2026	GCCCCGGGAGCACCGTGTGGGACCAAGTGTGTCAGTGTGCCGAGGGAACTTACCA	2C83
Qy	2024	TGGGACCCAGAACCTGCTGCATTTATGTGCCAATGGAAACTCTCCAAATAATGGGAAGGACA	2083
Db	2086	CGGCAGACGGGAGCATGGTGTGGCATGCCAGGGACCTCCAGGAGAGAAAGGGCA	2145
Qy	2084	AATGACTTGTGAAACCATGCCAAGAACCGGAAATTCTGGGCCCTGAGACCCAGAGC	2143
Db	2146	GCTCTCTGCGACCTTGCCTGGAG-----TGTGCCACGGCTCTTGGAG	2196
Qy	2144	TGCGAATATGTCTGAAATGTGAGGTCTGTGCAACCTGTGTGAAATAATGCGAGATGCGT	2223
Db	2197	CACCAACCTCACCGTGCGACGGTCACTGGCCACATGGCTGTT	2256
Qy	2204	TGCACCTTGGCAGCTCTGTGCCCTGGGACGTTTCAAGCTGAGCTGTTGCTGAACTCTG	2263
Db	2257	CAAGCCCTGTCAGCATGCCAATGGCACCTTCAAGGACCTGAGCAGGACCTATG	2316
Qy	2264	CTTCCCTGTGGAGGGCTTGGCACCAAATCATGGAGCTACTCTTCTGGACTCTG	2323
Db	2317	CTTCCCTGTGGGGCTCACCAACRAGCATGAGGGCCTTCAAGACTG	2376
Qy	2324	TGAAACCAGAGTCAATGTTCACTGGGACATTCTAACACCACTCACCTGGAACTT	2383
Db	2377	TGACACCAAACTGGTGTCCCAGGGACTACTAACACAGACATCCACGGCTATA	2436
Qy	2384	TGCTTGGCCAGTGGAACTACCCAGCTGAATTGGAAAATAATGTTGTTCTGCC	2443
Db	2437	TGCGTGTGCATGGCTCTATGGCTGACTGGCCGACTTCCGTCAGAACTTCTGCA	2496
Qy	2444	AGGAAATACTACGACTGACTTTGGCTTCAAAACATAACCCAGTAAAAACAGAAAG	2503
Db	2497	AGGAAACAAAGCACAGACTTGTGGCTTACCGTGTGGCTTACAGTGGCTCA	2556
Qy	2504	ATGTGAGGGAGCTGGGAGATTCACTGGTACATGTAATCCCACAACTCCAGCAA	2563
Db	2557	GTGIGGTGGAGTGGTGAATGGTCACTGGTATATGTCAGTCCCACACTCCGCAA	2616
Qy	2564	TTACCCAGCCAACCCGAGTCAACCCGACCATCAACCCCAAGGGCCGATCCT	2623
Db	2617	CTACCCAGCTGGTGGGGTGCATGGTGCATGGTGCATCAACCCCAACCCAGCGTAAGTCTC	2676
Qy	2624	GATCGTGGTCCCTGAGATTCTCCGCCATAGGGACGACTGTGGGACTACTCTGGTGT	2683
Db	2677	TATCGTGTACCAAGATCTCTCTGCAATCTGGGATAGTGGGAGCTCTCGTCAT	2736
Qy	2684	GCGGAADAACTCTCATCCAAATTCTGTGACAACATATGAAACCTTCCAGAACG	2743
Db	2737	GAGAAGAACTCATCCCACATCCATCTTACCACTTATGAGACTACAGGCG	2796
Qy	2744	CCCCATTCGCCCTCACCTCCAGGTTCAAGGTCCTACAGTCAATGAGACTACAGGACT	2803
Db	2797	TCCCATTTGCCCTCACCTCCGTTCACTGGAGACTCTGGTAACTTGTGAGCT	2856
Qy	2804	GAACAGCGTGTAGGGTCTCAAGGTCCTACAGTCAATGAGACTACAGGACT	2863
Db	2857	CAACAGGCCCTGGCTCCAGATTCCTATGTGAGCTTGTGAGCTATGAGCT	2916
Qy	2864	CATTGAAAGACATAGTTCGAGTGGCAGGGCTCTGATCTGAAACCCATCGAAATACT	2923
Db	2917	GGTAGAAGCATTTGCGGATCTGGTAAACCCACAGGAGTCTTGTGAAACCCAGAGATTT	2976
Qy	2924	TAAGGATAAGAAACTTATAAGGGCTCTGGTGTGATTCAAGTCCTGGCCATCCCCAGAACTATT	2983
Db	2977	AAAGACAGAAAGCTCATTAAGGCTCTTGTGAGCTTGTGAGCTATGAGCT	3036
Qy	2984	CAAGTACAGCCAGGACTCCAGAGTGTCAAGATCTTCTCGATTGTGAGCT	3043
Db	3037	CAAGTACAG-----AGAACACAGAGATGCTGCCAAAATCTTCATCAAGCTCCG	3093
Qy	3044	TTCACAAAGTGTGTCAGGTTTGTGAGACCTTACAAAT	3079
Db	3094	CTCCAAAGTTCAGGCTCTCTGAGCCCTAACAT	3128

Db	448	TGCCGGGCAAACGGGGCTGTCAAGCAGAGCTGTGAAACATGATGGGCAGCTATGAGTG	507	Qy	1529	-----TGTCAACCACTATGAGCTTCAAGAACATGATGGGCAGCTATGAGTG	1562
Qy	536	CTGCTGGCAAGGGGGTTTTCCTGACTGACATCACCACTCGTACCTGCTCCGGAA	595	Db	1546	CAACCTCCAAACACTGCACTGAGGCTGCACTGCTCCATTAAACAACTGCTCCCTAA	1605
Db	508	CCACTTGCGGGAGGGTTCTCTCCGGAAACCCGATTAACCTGTATCCGGCAGA	567	Qy	1562	GCTAAATGAGGGAAGTGTGAACTGTTGAA-----AAATGTTGAGTTGTTCCGGAGGCT	1615
Qy	596	AGGGGCTGAGTCATGATGAAAGGATCAGCTGGCTGTAGTCACATCTGCAAGGGCCC	655	Db	1636	GATCAGGATGCCAAATGGCTTGACCTGGAAACAAAGGAAAAGAGGGGCTGG	665
Db	568	AGGGGATGAAATGCTGAAACAGAACCCGCTGACCTGCAAGGAGACCC	627	Qy	1616	GGGACCGGCACTACAGGAAAGCACGGCTAGTAAAGAGACCTGGCTACGTTAACCT	1675
Qy	656	AAGGGGAGCTGCTGTAGTCAGGCTGCGCTGTAGTCAGGCTGCAAGAACAGAGA	715	Db	1666	CAGATCACGGGCAAGGGTGTGAGCTGAACTGCTGAACTCATTCATCCACCT	1725
Db	628	CAAGGGGTATGCCCTGTGAAAGAACCAAGGGAA	697	Qy	1676	TACATGAGCTCTGGCAAGAACAGGCTGGAGGAGGCTGGCGGACCCCTPAGGA	1735
Qy	716	CTGCTATCTGACTGTGAACTGGGACTCCCTGAGTGGATAACAGC	775	Db	1726	TAAGTGTGACTCTCTCGGAGGCAAAGGCCAAAGGGCCAGGCCAGGAGA	1785
Db	688	CTGTAATTGACATGCAACTATGTAACGGCGCTGCCAGCACACTGTGATGACAGA	747	Qy	1736	AATGTTA---TCACTGTTGAGTTGAGTTGAACTAACAAAGGGGTTGAGCTTC	1792
Qy	776	CGATGGCCCAAGTGCACCTGCACTCACAGTGCACAGATGGAGGAGCTG	835	Db	1786	GGTCACACAGGCTAACCTGGAACTGGGCAAGGGTCAAGGGAAAGAACAGCAG	845
Db	748	CGAGGGTCCCCTGGTGTGCTGCAATAAGTTGTTGCTCATACCGCAAGGACATG	807	Qy	1793	TTGTGACCTGAGGCTGATCTAAAGGAACTGAGGAGGCTCTGAAAGCCAATCCGGAC	1852
Qy	816	CTTGAGGGAGGACACTGTCTGGGGTGTGAAACACCACATCACTGGTGGAA	895	Db	1846	CTGEGGCTGCCCCCTGCTGACGCACTGAACTGAGGATCCCTGAAAGAT	1935
Db	808	CACTGGG-----	844	Qy	1853	GCTCGAGAAAGGCCCTCACAGGGAGGAGGCTTCAGGCTCACTCAGGCT	1912
Qy	896	TGGGATTAACGGGTGAAACGGGGCTCTCATGGAGCTGCTGCAAAATGGG	955	Db	1926	GCTCGAGAAACTCCATCAACAGGAACTCCCGAAATCTGAACGCCAGGA	1965
Db	845	-----CAACCGCTTCTTAATGAGACCTGCTGCTCAAAACGGGG	885	Qy	1913	CGTGGCTAAAGGCTCCCGAAATCTGAACGCCAGGA-----GAGTCCTGCG	963
Qy	956	CTGNGACCGACCTGTAAAGGATPACTTCGACGGTGTCACTGCTGTGGATT	10:5	Db	1966	GCTGGCCACAAAGGGGCTGGTAGCCGGGGGAGGAGGCGATGAGTCCTGAG	2025
Db	886	CTGTGACATTAAGGCTGCACTGAGTGGCTGCTGAGCTGGCCTGGCCT:	945	Qy	1964	AGTGGCCAGGGCAACCTGCACTTGTGCAAAATGGAACTTCCAGGAGGACA	2083
Qy	1016	CACTCTCCAGTTGGATGAAAGCATGATGAGTGGCTGCTGAAACCGCAATGG	1075	Db	2026	GGCCGGGCAACACCTGTGGCAACAGGAGCTGAGCTGGCAATGTTGTTACCA	2085
Db	946	CATCTGCAAGCCACAGAGACATGAAAGATGAGTGGCTGAAAGAAGG	1005	Qy	2024	TGGACGAAAGAACCTGCACTTGTGCAAAATGGAACTTCCAGGAGGACA	2136
Qy	1076	AGGTGTGATCACTTGTGAAAGCATGATGAGTGGCTGCTGAAAGAAGG	1135	Db	2086	GGGCCAGCCAGGAGCTGAGCTGCTGCACTGGGCACTGGGCACTTGGGAG	2145
Db	1006	GGGTGTGACACATATTGGCCACACAGTGGGCACTGGTGAAGAAGG	1065	Qy	2064	AATGACTGTGAACTGACATGCCAGAGCAAGGAAATTCTGGCCCTGTGAACCCAGAAC	2143
Qy	1136	ATTTAAATTAACTACAGTAGAGAAGTCTGCCAGATGAGTCCTTGTGATG	1195	Db	2146	GCTCCTGGCAGCCTTGGCTGGAG-----TGAGCCACAGGAACTGGGAG	2196
Db	1066	CTATAAGCTCATCAATGAGGAACTCCAGGATTAAGGATCTTGTGATG	1125	Qy	2144	TGGAAATATCTGTGAACTGAGCTGCTGCTGAACTTCTGAGTGTAT	2203
Qy	1196	GAACCTGGCAACAGCTGCACTACCACTGGGACATGCTGCTGGAAACCSAGG	1255	Db	2197	CACCAAGCTAACCTGGACCTGGCTGAGGTGAGCTGCTGAGTGGTT	2256
Db	1126	AACCTGTGACACATATGTGCAACACAGTGGGCACTTGTGAGTGGAAAGG	1:85	Qy	2204	TGCACCTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2263
Qy	1256	GTACACCTGTGACATGCTGCAACACTGGGAGCTATGCAACATACACGG	1315	Db	2257	CAAGCCCTGAGCCATGGGAGGAGCTGGGACCTGGGACCTGGGACCTGAT	2316
Db	1186	CTACCTGTGTTGATGGTATCAACCTGCTGCTGCTGCTGCTGCTGCTG	1245	Qy	2264	CTTCCCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2323
Qy	1316	AGGTGTGACAGACATGCTGCAACACTGGGAGCTATGCAACATACACGG	1375	Db	2317	CTTCCCTGTGGGGGGGGCTCACACAGGATGGCTGGGGGGGGGGGGGGGG	2376
Db	1246	AGGTGCCCTGGCTGATCAACTCTGGAGTGAATGAGCTGACTGTGCTG	1305	Qy	2324	TGAAACGAGCTGAGCTGCTGAAACATACCCGCTGAAATTGGTCTTCCTGCC	2443
Qy	1376	--GTACAGCTGCACTGTGAAATAAAAGACTGTGGAAGTGAAGGGCTCTGCCAC	1432	Db	2437	TGCTGTGGGGTCTTATGCCGACTCTGGCTGCTGCTGCTGCTGCTGCTG	2496
Db	1306	CCAGGGTCTGCTGCACTGCAATGCACTGCACTGCAACCTGTGCA	1365	Qy	2444	AGGAATACTAGCTGACTCTGGCTACCTCTGCAACACATACCCGAGTAA	2503
Qy	1493	CCTCTAGATGTCACTGTGCACTTGTGCAACATGCAAGGGCTGCAACCTGTG	1528	Db	2497	AGGRARCAAGCACAGACTGTGAGCTGCTGCTGCTGCTGCTGCTGCTG	2556
Db	1426	CTGTGGGACCCCCAGCCCCAGCCCCAGCCCCAGCCCCAGCCCCAGCCCC	1485	Qy	2504	ATGGGGAGGAGCTGGGAGATTGCTGAGATGGCTAACATGCCAGGAA	2563
Qy	1529	-----	1528	Db	2557	GTGTGGGAGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2616
Db	1486	CTGTGGGACCCCCAGCCCCAGCCCCAGCCCCAGCCCCAGCCCCAGCCCC	1545				

QY 2564 TTAACCGAAACACCGAGTGTACGTGGACATCAACCCAAAGCCCCAAGGCCGATCCT 2623
 Db 2617 CTACCACTGGTGGGTGATCTGAAATCAACCCAACCGAAAGTCCT 2676
 QY 2624 GATCGTGTCCTGTGAGATCTCTGGCCATATAGGGAGACTGGGACTACTGGTGT 2683
 Db 2677 TATCGTGTAAACAGATTCCTGGCATCTGGGATGAGTGTGGGACCTGCTCTCAT 2736
 QY 2684 GCGGAAGACCTCTCATCAATTCTGTGAAACATATGAACTGCAAGCTACCAAGC 2743
 DS 2737 GAGAAAGAACTCATCCATCTCCATTAACACTGTGAGCTGAGCTGAGCC 2796
 QY 2744 CCCGATCGCTTCACTTCAGSTCAAAGAAGCTGT3GATTCAGTTCAATGAAAG 2803
 Db 2797 TCCATTGGCTTACTGGCGTTCAAGGAACTGCTGAGTCAAGAAAGCAGGGC 2856
 QY 2804 GAAGAGCGCTAGGGTTCCAGGTCCTACATCTGACATAATGATGAGGACTACAGGAAC 2863
 Db 2857 CAACAGGCCGCTGGCTTCAGATTCCATGATGAGGACTATGAGCT 2916
 QY 2864 CATGAGAACATGTTCTGAGATGCAAGCTATGCTGACCTATGAACTTAACT 2923
 Db 2917 GGTAAGAAGCATGTGCAGATGCCGCTTAATGCCCTTAAGCCTGAAACCCGGAGTT 2976
 QY 2924 TAAGGATAAGAAACTTATCAAGGTTCTGTGATGCTGGCCATCCTCGAAACTTT 2983
 Db 2977 AAGGACAAAGCTCATCAAGGCTTCTGAGTGTGGCTAGTCACCCAGAACTACT 3036
 QY 2984 CAACTACACGCCAGGAGTCCCGAGAGATTTCCAAAGATGTTTATCCATTGCTACG 3043
 Db 3037 CAAGTACAG--AGAACACAGGAGATGCTGCCAAAATCTTCAATCAAGTGTCCG 3092
 QY 3044 TTCCAAAAGTCAGGTTTGAGACCTTAACAAAT 3078
 Db 3094 CTCCAAAAGTTCAGCTTCAGGCTCTGAGCCCTCAAAAT 3128

RESULT 15
 US-09-898-570-9
 Sequence 9, Application US/09898570
 Patent No. US20020123612A1
 GENERAL INFORMATION:
 APPLICANT: GERLACH, VALERIE L.
 APPLICANT: ELLERMAN, KAREN R.
 APPLICANT: MACDOUGALL, JOHN R.
 APPLICANT: SMITHSON, GLENDA
 TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USING THE SAME
 FILE REFERENCE: 15466-776CIP
 CURRENT APPLICATION NUMBER: US/09/898,570
 CURRENT FILING DATE: 2001-07-03
 PRIOR APPLICATION NUMBER: 60/198,293
 PRIOR FILING DATE: 2000-04-19
 PRIOR APPLICATION NUMBER: 60/198,645
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: 60/210,809
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/224,610
 PRIOR FILING DATE: 2000-05-09
 PRIOR APPLICATION NUMBER: 60/200,024
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/199,880
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/218,591
 PRIOR FILING DATE: 2000-07-17
 PRIOR APPLICATION NUMBER: 60/271,814
 PRIOR FILING DATE: 2001-02-27
 PRIOR APPLICATION NUMBER: 60/215,855
 PRIOR FILING DATE: 2000-07-03

PRIOR APPLICATION NUMBER: 09/639,446
 PRIOR FILING DATE: 2001-04-19
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 9
 LENGTH: 440
 TYPE: DNA
 ORGANISM: Unknown Organism
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: PCLYX
 OTHER INFORMATION: Zer7832_3 0.754
 JS-09-898-570-9

Query Match 25.2% Score 340.8; DB 1C; Length 444C;
 Best Local Similarity 62.1%; Pred. No. 9.7e-277; indels 192; caps 9;
 Matches 947; Conservative 0; X-matches 947;

QY 176 CGCGCGGTGCGGGCGTGCCTGGGGCGAGGATGAGATGAGTGTGGCCAGG 235
 Db 220 CACGCCCGCGCCCAATCACAAACCGC-GCAAGATGGCTGCTGGGG 178
 QY 236 GCTAGATGACTGGCCTGAGCCCTGTCAGAACACCCACCTCTAACAGTGTCT 295
 Db 279 GACTGACAACTGCACATGATCTATGCCAAACCCGGAGTAAATAGTGCAT 238
 QY 295 CTGCAAGCTGCTGAGCTGGAGGAGTGTAGGACATGGAGATGTGGAAA 355
 Db 239 CTGCAAGCTGCTGAGCTGGAGCTGGAGCTGGTGTGAGTGTGGAGCG 298
 QY 356 TGAGCTCATGGAGCTGCTCCATGAGCTGTGTTGAAATTCTGGCACT 415
 Db 299 AGGAGATAATGAGCTGGTGTGAGCTGGTGTGAAACTCTGGCAATTTGGGT 358
 QY 416 TTGTGTTGATGGCTCAGTTGGCTCATGAGCTGAGTGTGAGCTGGAGTGTG 475
 Db 353 CTGCTATGATGGATCCACCTGGCACATGAGCAACACTGCTGTTGGTGTGGAGTG 418
 QY 476 CCTGGAGAACATGGGCTGCCAGCATACCTGTGTAACGTCATGGGAGCTATGAGTG 535
 Db 419 TGCCGAGGCAACCGGGCTGAGCTGGTCACTGAGCTGGTGTGAGCTGGT 478
 QY 536 CTGCTGCAAGGGGGTTTCTGAGTGTGAGCTGGAGCTGGTGTGAGCTGGAG 595
 Db 479 CCACTGCCGSGAAGCTTCTCCAGACACCAGATACTGTTGGTGTGGAG 538
 QY 596 AGAGGGCTGAGCTGATGAAATAGGAACTGGCTGAGTGTGAGCTGGAGGCC 655
 Db 539 AGAGGAGTAATGCTGAGAACAGAAACCCAGGAGACAGAC 598
 QY 656 AACGGGCAAGGTCCCTTGAGTCAGGCTGAGCTGGCTGTTTGGAGCTGGAG 715
 Db 599 CAAAGGGGGTATGGCTGAGTCAGGCTGAGCTGGCTGAGCTGGAG 658
 QY 716 CTGCTCATGACTGACCTGAACTGAACTGGCTGAGCTGGAGCTGGAG 775
 Db 659 CTGTAATGACATGCAACTATGTAAGGCTGCTGAGCTGAGCTGGAGCAAG 717
 QY 776 CGATGGCCGAGATGAGCTGGCATCCACATGAACTGAGCTGGAGCTGGAG 835
 Db 718 ----- 717
 QY 836 CCTGAGGAGGAGGAGGAGCTGGTGTGGAGCTGAGAGGAAACACACATGCTGGTGA 895
 Db 718 ----- 717
 QY 896 TGGGATAAACGGGTTGAAACGGGGCTGCTGTCATGGAAACGTTGCTGCTGGGG 955
 Db 718 ----- -AGGGAGACTCTGCTGCTGAGCTGGGG 746
 QY 956 CTGTCACGCGACCTGAGGATACTCCACAGS3GTGCTCAGTGTCTGCTGGATT 1015
 Db 747 CTGTCACGAGTGGCTGAGCTGAGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTG 826

Qy	1616	CACTCTCCAGTTGGATGGAAAGACATGTAAAGATATTGATGAGTGGTCCAGACCCGCAATGG	1075	D5	1887	GGAGAGAGAAGGGCAGTCCTCTGCACCCCTGGCGCTG	-----GATGTGATGCCCA	1937
Db	807	CATGTCGACGCCAGACAGGAGACCTGCAAGAGATATAGATGAGTGGCCGGTTAACACCG	986	CY	2129	GAAGACCCAGAAGCTTGGATATGCTGTTGAATGTTGGAGCTGTGCTGAAATA	-----CAACCTGGTGAAATA	2188
Qy	1676	AGGTGTGATCATTTCTGCAAATAATCGTGGCGAGTTTGACTGGGTGAGAAAGG	1135	Db	1938	CGGCCCTTTGGGCCACCAAGTCACACGTC	-----ACCTGGCAACCTGGCAACA	1997
Db	867	GGGTGTGACCATATTGGCCAAACACAGTGGCACGTTGAAARGCAGTGGAAAGG	926	QY	2189	TCTGCAATGGCTTGGACCTTGGCCAGCTGCTGCGCTGAAAGC	-----CTCCGGCTGAAAGC	2248
Qy	1136	ATTAATTTAAAGATAGAAACTCTTCCCACATGTTGAGTGGATGCTTGGATAG	1195	Db	1938	CTCTGTAGATGGTTCAAGCCCTGCAAGCTTGGCCACCTAACCTGTCAG	-----GATGTCCTGAGC	2057
Db	927	CTTAAGCTCTTCATCAATGAGAAACTCCAGATAAGACGAGTGTCTTGTGATG	986	QY	2249	TGTGGAACCTCTGTTCCCTGTGAGGAGGCTGACCAAAAGATGGGAGGCTAC	-----TGTGAGGAGGCTAC	2308
Qy	1196	GACCTGTGACCAACAGCTGATCAACCCTGGACATTGCTGTTGACCGAGG	1255	Db	2038	AGGAGCAGGCCCTATGCTCCCTGTGTTGGGGCCCTACACCCACAGATGAAAGG	-----GGGGCCAT	2117
Db	987	ACCTGTGACCAACATGTCAGTCTCGCCATCGTGG	1146	QY	2303	TTCCTTCTAGGACTGTGAAACCAAGAGTTCAATGTTCACTGSGACATTCT	-----TACACACACAC	2368
Qy	1256	GTACACCCGTATGGCTTCAACCACTGTGGAGCAGCATCAACCG	1115	Db	2118	TTCCTTCAAGACTGTCAGTGTGACCAAAAGTCAGTCAGGCACTACAAAC	-----CAGTGTGAACTAA	2177
Db	1047	CTACCTGTGTATGTTGATCATCCACTGTCGGATGTGGATGAACTACCGGGG	1106	QY	2369	CACTCACCGATGTTATGGCCAGTGGGACATACAGCTGTAATTGAA	-----GATGAAATATAA	2428
Qy	1316	AGGTGTCAAGAGGTTGTGAAACAGTGGCGAGCTATACTCCAGTGGCTG	1375	Db	2178	CATCCACCGCTGTTATGGCCCTATAGGCCGACTTCCGGCT	-----TACAGTGTGAACTTT	2237
Db	1107	AGGTGCGCTTGGCATCTACCTGGACCTACAGCTGATCTGGGAGTGTGG	1166	QY	2429	TGTGTTCTGTGCCCCAGAAATACTACGAGCTACTGATGCTCCACAAAC	-----ATAACCCA	2488
Qy	1376	--GTACAAGCTCCACTGGATAAAAAGACTGTGTTGGAAAGGTGCTGCCAAC	1432	Db	2238	CTGAGCCGCTGTCAGGAAACAAAGACAGACTGTGATGCTCTGAGCT	-----TGTGAGTGTGGCCA	2297
Db	1167	CGGGTCGGTGACTGGATGGAAAGATGTGACAGACGACTGAGGTGAGG	1226	QY	2489	GTGTAAAAACAGAGAATGGAGGGAGCTGGGAGATTCACTGGTAACTG	-----GGTAACTTC	2548
Qy	1433	AGTGTGTCACCCGGTGTGICCTGCACTGGTAAGAGTGTGGAGAGCAGGTGTT	1492	Db	2298	ATGGAAGATGTCAGTGTGGGGAGCTGGTGTGAACTTGTGAACTG	-----TGTGAACTGTC	2357
Db	1227	TCTCTGGGGCTCGAAAGCCATGGCTACGCTTGGCATCAACGGACACCTGTC	1296	QY	2549	AAACTACCGGAAATTACCCGGCAACAGGATSTACGGGACCATCAAC	-----CCCAAC	2608
Qy	1493	CTGTACATGTAATCTGGTCACTGGTACCTGGTCACTGGTCACTGGTCA	1546	Db	2358	CAACTACGGGAAACTACCCGGCTGAGTGTGAACTTCACGGAA	-----ACCCCAAC	2417
Db	1287	CCTGACTCTGTCCTCCAGGGCGATTTTCCAGGGTCACTGGTCACTAAACA	1346	QY	2609	CAAGCGGCACTCTGATGTTGCTGCCATAGGGAGACTGTGG	-----TCTGGTAACTG	2668
Qy	1547	AGTGTAACTTTAACTGAAGCCAAGCTGACTTTGA-----AAATGCTGAGCT	1600	Db	2418	CAAGCGCAAGATCTTACGTGTTGATGAGATCTTGTGAGTGTGG	-----TGTGAGTGTGG	2477
Db	1347	ACGGGCTCTTCATCAGATCAGGATGCCAACTGGTCACTGGTCAACAGGCA	1406	QY	2669	GGACTATCTGGTATGGCAAGGAAACCTCTCATCCAAATTCTGTGAA	-----ACATATGAAAC	2728
Qy	1601	GTTCCTGGAGGGCTTCAGGGCTTCAGGATGCCAACTGGTCACTGGTCA	1660	Db	2478	GGAGTCTCTGTCATGAGAAAGACTCATCCCATCTTGTGAACTG	-----TGTGAACTG	2537
Db	1407	ACAGAGGAGCTGGGGAAACCAAGGGGGCTGGTGGACCTGGTCACTGGTCA	1466	QY	2729	CCGACCCAGGAAAGCCCATGGCTCACTCCAGSTCAAGAAAGTGTGATT	-----GTTGTT	2788
Qy	1661	CGGTCTGTAACCTTACATCGAGCTGGAGAAGTCCAGGCTCTGGCGAAC	1120	Db	2538	CCGACCCATGAGGGAAAGCAGGGCTAGGGTTCCAGGTTCTGGTCA	-----CTGGATCAGT	2597
Db	1467	CACCTTCATCCACCTTAAGTGTGACTCTCTCGGAGGGCAAGGGCGGCGAC	1526	QY	2789	CAAGTCCATGAGGGAAAGCAGGGCTAGGGTTCCAGGTTCTGGTCA	-----GATATGTTGA	2848
Qy	1721	AGACCCCTTAAGGAAATGTTA--TCACGTTGAGTTGAACTAACCAAAA	1177	Db	2538	CAAGACAGCGGGCCACAGGGCCACAGGGCCCTGGCTCCAACTTAC	-----CTATGATGAA	2657
Db	1527	CCCTCCAGGAAAGGGTCACTGGTCACTGGTCACTGGTCACTGGTCA	1586	QY	2849	CCATCAGGAAATACTTAAAGGATAAGAACCTTACAGGCTGTTGATG	-----TGTGCTGGCCA	2968
Qy	1778	CGAGSTGACACCTCTGTGACCTGAGCTGATCTGGAAACCGGAAAGCGCGCTCG	1837	Db	2718	CCACCAAGGAGATTTAAAGGACAGAACCTCTGCTTCAASGCCCTAG	-----CTGGTCA	2777
Db	1587	AGAAACACAGCAGCAGCTGGTGTGCTCTCCACAGGAAATGGAGAACGG	1646	QY	2959	TCCCGAGACTATTCACTGAGTGTGCTGAGATGTTCCAGATGCT	-----TGTGCTGGCT	3028
Qy	1838	TAAGGCTTACCCGCTCACAAAGGCCGTCACAGGGAGAGTTCACCTCC	1897	Db	2778	CCCCAGGAACTCTGAGTGTGCTGAGTGTGCTGCAAAATCTT	-----GGAAATCTT	2834
Db	1647	AGGATCCTGTGAGATGTCAGTCACTAACGGAGCCGCTCTGGCGCTGGC	1706	QY	3029	CATCGATGTTGCTAGTCAAGTGGCTGAGATGTTCCAGATGCT	-----TGTGCTGGCT	3078
Qy	1898	AGGCACTAACCTCGACCTGGCTGGTAAAGGCTCCACCTGAAACATCTG	1952	Db	2835	CATCAAGGTCTCTAGGCTGTCATGCTCAAGTTCAGGCTCCTAC	-----AAATGAACT	2884
Db	1707	AGGCCCTGATATGGTGGCCACAGGGCTGGTAAAGGGAGAGGAGAGC	1786	Qy	3099	TTGGACTTATATGATGGAGGAGCTGGAGCTGGTCACTGGTCA	-----AAATGGAACTT	2128
Qy	1953	---GAGTCCSTGGGACTGGCCAGGCTCATGCAAGAAAACCAATGTCAG	2008	Db	1827	GGAACCTTACACGGCCAGCGAGCTGGTGTGCGCCATGCCGACCTTCCA	-----1896	1896
Db	1767	GATGGAGCTCTTAGGCCGGCAGACGGCTGTCATGCTGAGTGTGGCT	1826	Qy	2069	AAATGAGGAAGGACAATGTCAGACCATGCCCAGGAAATTCTGGGGCCCT	-----AAACCCGAA	2128

Search completed: October 22, 2003, 22:02:39
Job time : 969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK: nucleic - nucleic search, using sw model

Run on: October 22, 2003, 17:27:46 ; Search time 238 Seconds

(without alignments); 6930.457 Million cell updates/sec

Title: US-09-747-371-1

Perfect score: 3737

Sequence: 1 ggctatcgccacacctccc.....ctaaggacttctggagacat 3737

Scoring table: IDENTITY_NJC

Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *

1: /cgn2_6_ptodata/2/ina/5A_COMB.seq;*

2: /cgn2_6_ptodata/2/ina/5B_COMB.seq;*

3: /cgn2_6_ptodata/2/ina/6A_COMB.seq;*

4: /cgn2_6_ptodata/2/ina/6B_COMB.seq;*

5: /cgn2_6_ptodata/2/ina/PC703_COMB.seq;*

6: /cgn2_6_ptodata/2/ina/backfilesi.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query Score Match Length DB ID Description

Result No.	Score	Match	Length	DB ID	Description
C 1	257	6.9	291	4 US-09-389-681-246	Sequence 246, App
C 2	257	6.9	291	4 US-09-620-405B-246	Sequence 246, App
C 3	257	6.9	291	4 US-09-339-338-246	Sequence 246, App
C 4	257	6.9	291	4 US-09-433-826B-246	Sequence 246, App
C 5	257	6.9	291	4 US-09-604-28-A-246	Sequence 246, App
C 6	213.4	5.7	241	4 US-09-439-313-412	Sequence 412, App
C 7	213.4	5.7	241	4 US-09-352-616A-412	Sequence 412, App
B 8	83	2.1	3373	2 US-08-897-443-2	Sequence 1, Appli
B 9	78.6	2.1	1531	2 US-08-833-963C-1	Sequence 1, Appli
B 10	78.6	2.1	2018	3 US-08-980-514-2	Sequence 2, Appli
B 11	76.6	2.0	1254	4 US-09-312-283C-67	Sequence 67, Appli
B 12	75	2.0	1260	3 US-09-188-930-67	Sequence 67, Appli
B 13	71.8	1.9	2550	2 US-08-884-072-2	Sequence 2, Appli
B 14	71.8	1.9	2550	4 US-09-11-168-2	Sequence 2, Appli
B 15	70.2	1.9	2461	1 US-08-282-141-1	Sequence 1, Appli
B 16	63.6	1.7	1018	3 US-09-188-930-259	Sequence 259, App
B 17	63.6	1.7	1018	3 US-09-312-283C-259	Sequence 259, App
B 18	61.4	1.6	3546	3 US-08-872-757-3	Sequence 3, Appli
B 19	61.4	1.6	3546	4 US-09-850-048B-3	Sequence 3, Appli
B 20	59.4	1.6	2512	4 US-09-248-57-1	Sequence 1, Appli
B 21	58.5	1.6	3284	6 5258988-3-	Sequence 1, Appli
B 22	58.8	1.6	3290	1 US-07-985-691-1	Sequence 1, Appli
B 23	58.8	1.6	3290	1 US-08-436-804-1	Sequence 1, Appli
B 24	58.8	1.6	3290	1 US-08-267-387-1	Sequence 1, Appli
C 25	58.8	1.6	6811	3 US-08-651-472-67	Sequence 67, Appli
C 26	58.8	1.6	6811	3 US-08-358-928-67	Sequence 67, Appli
B 27	56.4	1.5	2397	6 5258288-2	Patent No. 5258288

ALIGNMENTS

RESULT 1
US-09-389-681-246/C
Sequence 246, Application US/09389691A
Patent No. 651237
GENERAL INFORMATION
APPLICANT: Yuqiu, Jiang
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: DIAGNOSTICS FOR THE TREATMENT AND
METHODS FOR THEIR USE
FILE REFERENCE: 21011470C3
CURRENT APPLICATION NUMBER: US/03/389,681A
NUMBER OF SEQ ID NOS: 463
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NC 246
LENGTH: 291
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: [1]-[291]
OTHER INFORMATION: n=A,T,C or G
US-09-389-681-246

Query Match: 6.9% ; Pred. No. 2.2e-5 ; DB 4 ; length 291;
Best Local Similarity: 97.5% ; X-matches: 6 ; Indels: 1 ; Gaps: 1;
Matches: 268 ; Conservative

Qy 2006 GCTGGGACCTATTATGTTGAGGCCAGAAAGCC-TGCAATTATGGAACT
Db 275 GGTGGGACCTATTATGGAGCAGAAAGCTTGCATTAATGCAATTATGGAACT
Ct 216

Qy 2065 TCCAAAATGGAGGAAATGACTTGAACTTGCACCATGCCAACGAAATTCTGGG
Db 155 CCTCTGAAACCCAGAAAGCTTGAATAATGCTGAAATGGAGCTGTC
Ct 214

Qy 2185 AATATTGCGAGTGGTTGACCTTGCACCTCTGCTGGCACCTCGCTGG
Db 215 TCCAAAATGGAGGAAATACTTGAACTTGCACCATGCCAACGAAATTCTGG
Ct 156

Qy 2185 AAATTCTGGCAGANGCTTGACCTTGCAGCTTGCCTGGCACCTCGCTGG
Db 95 AAATTCTGGCAGANGCTTGACCTTGCAGCTTGCCTGGCACCTCGCTGG
Ct 216

Qy 2245 AACCTGGTGAACCTCCCTGCTGGGACCTCGCTGGCACCTCGCTGG
Db 2279

Qy 2245 AACCTGGTGAACCTCCCTGCTGGGACCTCGCTGGCACCTCGCTGG
Db 35 AACCTGGTGAACCTCCCTGCTGGGACCTCGCTGGCACCTCGCTGG
Ct 1

Best Local Similarity 98.7%; Pred. No. 4.5e-46; Matches 225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 2350 GACATTTCTAACACCAACTACCGATGATTCTGTGCCAGTGGAAATACCCAGC 2409
Db 15 GACATTTCTAACACCACTACCGATGATTCTGTGCCAGTGGAAATACCCAGC 73

Qy 2410 CTGAATTGGAAAAATAATTGTGTTCTGCCAGGAAATACCGACTGACTTGTG 2469
Db 74 CTGAATTGGAAAAATAATTGTGTTCTGCCAGGAAATACCGACTGACTTGTG 133

Qy 2470 GCTCCACAACATAACCCAGTGTAAAACGAAATGATGTCGAGGAGCTGGAGATTCA 2529
Db 134 GCTCCACAACATAACCCAGTGTAAAACGAAATGATGTCGAGGAGCTGGAGATTCA 193

Qy 2530 CTGGGTACATGTAACCTACCCAGTGTAAAACGAAATACCCAGCAAA 2577
Db 194 CTGGGTACATGTAACCTACCCAGTGTAAAACGAAATACCCAGCAAA 241

RESULT 8

US-08-897-443-2
Sequence 2, Application US/38897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hallman, Jennifer L.
Corley, Preti Lal, Preti
Corley, Neil C.
Srah, Parvi
Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Disquette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3373 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UTRSNCT02
CLONE: 681719

US -08-897-443-2

Query: Match 2.2%; Score: 83; DB 2; Length 3373;
Best Local Similarity 50.7%; Pred. No. 1.6e-11;
Matches 253; Conservative 0; Mismatches 240; Indels 6; Gaps 2;

Qy 900 TCCAGGATCTGTGCCATTGGAGCACCAACTGTGAGCAGCTCTGTGAAATGTGCGG 949
Qy 905 GAGGTGTCACTGGAGTTCCTGTTGATTCCTCTCCAGTGGGAGACATGTA 1044
Db 950 GCTCTCTCTCTGCAGTCTACAGTGCCTACGCCCTGGCTGAGGAGGGTG 1009

Qy 1045 AAGATATTGATGACTGCCGACCCGCAATGAGGTGTCATTTGCAAACATCG 1104
Db 1050 TGGCTGTGACTACTGTGCTCGAARACACCGATGTGACATGATGTC 1069

Qy 1105 TGGGGATTTGACTGCCGTCGAGAAGGATTAAATTAACAGATGAAAGCTCT 1164
Db 1070 ATGGCTCCACCTGCGATGAGGTTCACCCGATTAACCAAAACGT 1129

Qy 1165 GCCAGATGGATGAGCTGACCCGCTGACCTGTAT--GCCTTCACCC 1221
Db 1130 GCACAAAGTAGACTTGCTCATCTAAATCACGGATGTCGAGCTGTTACA 1189

Qy 1222 ACCCTGGAAATTGCTTGTGCACTGGAGGTAACCCGTAT--GCCTTCACCC 1278
Db 1190 CAGATGATTCCTATCCTCCACTGCTGAAAGCTTAACTCTGAAATGAAA 1249

Qy 1279 ACTGTGGAGACACCAATGAGTGGAGCATAACACGGAGCTSTACGGAGCT 1328
Db 1250 CCTGGAGAGAGTAACTCTGACTGCTGACTGCTGAGTGGCTGCGTC 1309

Qy 1339 ACACAGTGGCAGCTATGATGCTGAGTCACAGCTCCACTGGAAATAAA 1398
Db 1313 ACATGGAGGAGTACTACTGGCTGGCACCGTGGTACACTTGACCCATGGCA 1369

Qy 1393 AGAGCTGTGGAAATGAA 1417
Db 1370 AAACCTGCAACCGAGCTGGAA 1388

RESULT 9

US-08-833-963C-1
Sequence 1, Application US/08833963C
GENERAL INFORMATION:
PATENT NO.: 5916769
APPLICANT: Olseni, et al.
TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Ave
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patenter Release #1.0, Version #1.3.C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,963C
FILING DATE: 11-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Broces, A. Anders
REGISTRATION NUMBER: 36,723
REFERENCE/DOCKET NUMBER: P258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-329-8534
TELEFAX: 312-329-8439
INFORMATION FOR SEQ ID NC: 1:
SEQUENCE CHARACTERISTICS:

RESULT 10
 Sequence 2, Application US/08980514
 Patent No. 6,004,753
 GENERAL INFORMATION:
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Shan, Purvi
 TITLE OF INVENTION: HUMAN S1-S-ECMP-LIKE PROT.
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/980,514
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:

LENGTH: 1531 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 125..1453
 US-08-983-963C1
 Query Match 2.1%; Score 78.6; DB 2; Length 1531;
 Best Local Similarity 54.6%; Pred. No. 1..5e-10;
 Matches 206; Conservative 0; Mismatches 159; Indels 12; Gaps 2;
 Qy 210 GAGGATCTAGTGAAGTGCCAAAGGGTAGATGACTGCCATGGCGACCCCTGTGTCAG 2.9
 Db 488 GTGGATGTCAGAGTGCCAGCCCTGACGACTGCAAGGAGCTGCATGGCT 54.7
 Qy 270 AACACACCCACTCTAACAGTGTCTCTGCAAGCTGGCTAACCAAGGGGAGGGAG 329
 Db 548 ACTGCGCTGGCTCTATCGACCTGCGACCTGCTGCTGCTGCTGCTGCTG 607
 Qy 330 TGTGAGGACATCGTGAATGTGGAATGACTCATGGAGCTGTGTCAATGACTGTG 389
 Db 608 TGTGAGGACATAGGAGTGTGCGAGCCCTGACGACTGCGCTG 658
 Qy 390 ATATACTCCGCAAATTATCGTTGACTTGTGTTGACTTGCTCATGACGT 449
 Db 659 AACCTGGCTGGCTCCATGACCTGGCTGGCTGGCTTAACAC 71.8
 Qy 450 CATAATTGCTCTGATGGAGCTGCGCTCGAGACATACCTGT 529
 Db 719 CGCTCTGTTGACTTGTGATGGAGCTGCGCTCGAGACATACCTGT 775
 Qy 510 GTCAACGTCATGGGGCTPATGAGTGTGCTGCTGCTGCTGACTGACAT 569
 Db 776 TTCAACTCCPAGGGACCTCTGTGCTGCCACAGGCTATGAGTCATGGGT 835
 Qy 570 CAGCAACCTGGCTTA 566
 Db 836 GGCTTCTCCCTGACAGCA 852
 RESULT 11
 Sequence 67, Application US/09312283C
 Patent No. 6,573,955
 GENERAL INFORMATION:
 APPLICANT: Watson, James C.
 APPLICANT: Strachan, Lorna
 APPLICANT: Sleeman, Matthew
 APPLICANT: Onrust, René
 APPLICANT: Murison, James G.
 APPLICANT: Kumble, Krishanand D.
 TITLE OF INVENTION: Compositions isolated from Skin Cells
 FILE REFERENCE: 1100C.1011C2
 CURRENT APPLICATION NUMBER: US/09/312,283C
 CURRENT FILING DATE: 1999-05-14
 NUMBER OF SEQ ID NOS: 425
 SOFTWARE: Fast-SEQ for Windows Version 4.0
 SEQ ID NC 67
 LENGTH: 1254
 TYPE: DNA
 ORGANISM: Mouse
 US-09-312-283C-67

Query Match 2.0%; Score 76.6; DB 4; Length 1254;
 Best Local Similarity 51.0%; Pred. No. 4.5e-10;
 Matches 188; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

Qy 1068 CGCAATGGGGTTGTGATCATTTCTGCAAACATCGTGGGAGTTTGACTCGGGTC 1:127
 Db 186 GCGPATGGTACTGCAGGAGCTGTGCAATTCCTTACAGT 2:39

Qy 1128 AGAAGAGGATTAAATTAAACAGTGAAGCTTGCAAGATGGTGAATGGCTCT 1:87
 Db 240 ACCCTGGTCAACCTCAACGATGATGGAAAGGTCTGCAGATGTGAAACAGTGAA 2:59

Qy 1188 TTGGATAAGGACCTGACCACAGTGCATCAACCCCTGGCAGATTGCTTGCTGC 1:247
 Db 300 ACTGAAACCCCTGTTGACCTGCGTCAACACTATGGTTTATCTGCCGTGT 3:59

Qy 1248 AACCCAGGTACACCT--GATGGCTTCAACCACTGGAGCACCAATGGTGCAGC 1:304
 Db 360 GACCTAGGATTAAGCTGGAGGAATGGTCAACTGGAGTATGGTGCAGC 4:19

Qy 1305 ATCACAAACGGGTGAGGTCTGGAGGTCTGGAAACAGTGGCAGCTATGGCSAG 1:364
 Db 420 TTCTGGAGTTCCTGTGCACTGAGTGTGAAACCCGGGCTCATCTCTGCTCA 4:79

Qy 1365 TGCCACCCCTGGTACAAGCTCAGTGGAAATAAAAAGACTGTGGAACTGAAGC 1:19
 Db 480 TGCCCTAACGCTACGTCTGGAAAGATAACCGAAGTGTCCAGGATATCAATG 5:34

RESULT 13
 Sequence 67, Application US/C8884C72
 GENERAL INFORMATION:
 APPLICANT: Bandman, Oiga
 APPLICANT: Corley, Neil C.
 APPLICANT: Giegerl, Karl J.
 TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatibl
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/C8884C72
 FILING DATE: Herewith
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-C333 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2550 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: CCNNCTC1
 CLONE: 45517

US-09-188-930-67
 Sequence 67, Application US/C9188930A
 GENERAL INFORMATION:
 APPLICANT: Watson, James D.
 TITLE OF INVENTION: Compositions Isolated From Skin Cells
 TITLE OF INVENTION: and Methods For Their Use
 FILE REFERENCE: 11000.1.01.c1
 CURRENT APPLICATION NUMBER: US/09/188,930A
 CURRENT FILING DATE: 1998-11-09
 NUMBER OF SEQ ID NOS: 348
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 67
 LENGTH: 1260
 TYPE: DNA
 ORGANISM: Rat
 US-09-188-930-67

Query Match 2.0%; Score 75; DB 3; Length 1260;
 Best Local Similarity 51.8%; Pred. No. 1.2e-09;
 Matches 184; Conservative 4; Mismatches 164; Indels 3; Gaps 1;

Qy 1068 CGCAATGGGGTTGTGATCATTTCTGCAAACATCGTGGGAGTTTGACTCGGGTC 1:127
 Db 246 AACCCCTGGCTTACCCCTCAAAGATGTGGAAAGGTCTGCACACTTGCTGTACGFT 2:45

Qy 1128 AGAAGAGGATTAAATTAAACAGTGAAGCTTGCAAGATGTGATGAGTGTCT 1:187
 Db 240 ACTGAAACCCCTGGTCAACCTGAGCTCAACACCTGTGCAAGTGAAGTGTCT 3:65

Qy 1248 AACCGGGTACACCT--STATGGCTTCAACCACTGTGGAGCACCAATGGTGC 1:104
 Db 366 GACCCARGATGAACTGGGAAAGATGGCAATTCACTGCATGAGTGTGARC 4:25

Qy 1305 ATCACAAACGGGGTTGTGATCATTTCTGCAAACATCGTGGGAGTTTGACTCGGGTC 1:364
 Db 426 TTCTGGAGTTCCTGGAAAGGTCTGCACACTTGCTGTACGFT 2:45

Qy 1365 TGGATAAGCTGGAGCTGAACTGGTGGAAATTTGCTTGCTGC 1:247
 Db 480 ACCGAAACCCCTGCTGAAACCTGCAACACTATGGTGTGAACTGATGTC 1:277

Qy 1428 AGAAGAGGATTAAATTAAACAGTGAAGCTTGCAAGATGTGATGAGTGTCT 1:187
 Db 240 ACTGAAACCCCTGGTCAACCTGAGCTCAACACCTGTGCAAGTGAAGTGTCT 3:65

Qy 1488 TTGGATAGGACCTGTGAGCTGACCTGAGCTGACATTTGCTGTACGFT 2:45

Db 306 ACTGAAACCCCTGGTCAACCTGAGCTCAACACCTATGGTGTGAACTGATGTC 1:277

Qy 1548 AACCGGGTACACCT--STATGGCTTCAACCACTGTGGAGCACCAATGGTGC 1:104
 Db 366 GACCCARGATGAACTGGGAAAGATGGCAATTCACTGCATGAGTGTGARC 4:25

Qy 1608 CGCAATGGGGTTGTGATCATTTCTGCAAACATCGTGGGAGTTTGACTCGGGTC 1:127
 Db 898 CGCTATGGTTAGTGCAGACGCTGTGCAATGTCGATCTGATCATGTC 9:57

Qy 1668 AGAAGAGGATTAAATTAAACAGTGAAGCTTGCAAGATGTGATGAGTGTCT 1:187
 Db 958 AACCCCTGGTTACCTCACTGAGATGGAGGTCTGGCAAGTGTCTGTC 1:217

Qy 1728 AGAAGAGGATTAAATTAAACAGTGAAGCTTGCAAGATGTGATGAGTGTCT 1:187
 Db 1886 TGGATAGGACCTGTGAGCTGACCTGAGCTGACATTTGCTGTACGFT 2:45

QY 1186 CTTGGATAGGACCTGTGACCACAGCTGCATCAACCACCCCTGGCACATTGGCTTGTGCTT 1245
Db 86C -TC TGAGGCCGCTGAGCAGTC TGCCTGAACTCCCAGGAGTACA CCTGCCACT 918
QY 1246 GCAAACCG 1252
Db 919 GTGAGGG 925

Search completed: October 22, 2003, 18:45:04
Job time : 241 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 11:24:43 ; Search time 47 Seconds
 (without alignments)
 2044.096 Million cell updates/sec

Title: US-09-747-371-2
 Perfect score: 5605
 Sequence: 1 MGYAGRNRPGAAWAVLILLL.....PRSFIRLRSKVSFLRPLYK 999

Scoring table: BLOSUM62
 GapOp 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-76.*
 1: pir1;
 2: pir2;
 3: pir3;
 4: pir4;*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	863.5	15.4	1620	2 T2283	hypothetical protein MEGF6 Protein - rat
2	841.5	15.0	1574	2 T11954	fibrillin-2 precursor
3	727	13.0	2918	2 A54105	fibrillin-2 precursor
4	724	12.9	2907	2 A52728	fibrillin-1 precursor
5	691	12.3	2871	2 A55367	fibrillin-1 bovin
6	691	12.3	3002	2 A47221	fibrillin-1 precursor
7	690	12.3	2871	2 A55224	fibrillin-1 precursor
8	687	12.3	741	2 T46488	hypothetical protein
9	674.5	11.1	1394	2 A356226	transforming growth
10	592	10.6	1742	2 A38261	masking protein
11	577.5	10.3	1221	2 A49577	latent transforming
12	570	10.2	1820	2 A51494	fibulin, splice isoform
13	567	10.1	685	2 S78940	fibulin-1 precursor
14	565	10.1	683	2 C36346	fibulin-2 precursor
15	560	10.0	1184	2 A5184	fibulin, splice isoform
16	558	10.0	705	2 S34968	fibulin-1 precursor
17	553	9.9	601	2 B36346	fibulin-1 splice isoform
18	552.5	9.9	2319	2 T42930	latent transforming
19	540.5	9.6	1251	2 A52723	hypothetical protein
20	530	9.5	1372	2 T25333	hypothetical protein
21	529.5	9.4	798	2 T22293	notch3 protein - h
22	526.5	9.4	2321	2 S78549	notch3 protein - h
23	524.5	9.4	2318	2 S45306	notch3 protein - h
24	514.5	9.2	689	2 T42760	fibulin, splice isoform
25	509.5	9.1	589	2 T43210	fibulin-1D precursors
26	509.5	9.1	2703	1 A24420	notch protein - fr
27	507.5	9.1	3507	2 T34513	hypothetical protein
28	502.5	9.0	2524	2 A35844	notch protein - Af
29	497	8.9	2555	2 A40043	notch protein homo

3C	494.5	8.3	2531	2 T31070	notch homolog - se
31	493	8.3	1263	2 A49175	Motch B protein -
32	489	8.7	2531	2 S-8198	notch-1 protein homo
33	489	8.7	2531	2 A46019	notch-1 protein -
34	487.5	8.7	1964	2 T05059	notch4 - mouse
35	487	8.7	2352	2 T30231	Notch homolog prot
36	484	8.6	2437	2 S42612	transmembrane prot
37	479	8.5	2471	2 A49128	cell-fate determin
38	466	8.3	1064	2 A40336	fibropellin Ia - s
39	443.5	7.9	1429	2 S06344	homeotic protein -
40	441.5	7.9	1274	2 T42017	cysteine rich prot
41	43.8	7.8	2217	1 EG5MG	epidermal growth f
42	431.5	7.7	1133	1 EGRC	hypothetical prote
43	423	7.5	1827	2 T34288	intrinsic factor B
44	422	7.5	3623	2 T38816	epidermal growth f
45	421	7.5	1207	2 EGHC	

ALIGNMENTS

RESULT 1					
T22783	hypothetical protein Y64G1CA.f	- Caenorhabditis elegans			
C;Species: Caenorhabditis elegans					
C;Date: 15-Oct-1999 #sequence_change :5-Oct-1999 #text_change :5-Oct-1999					
C;Accession: T227283					
R;Anscombe, R.					
Submitted to the EMBL Data Library, September 1999					
A;Reference number: 220336					
A;Accession: T22783					
A;Status: Preliminary; translated from GB/EMBL/DDBJ					
A;Molecule type: DNA					
A;Residues: 1..622 <NTB>					
A;Cross-references: EMBL:AU110498; NID:el542303; PDB:CB54471.; CESP:Y64G1CA.f					
C;Genetics:					
A;Gene: CESP-Y64G1CA.f					
A;Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 549/1; 546/1; 545/1; 601/1; 625/1					

Query Match	Score	Best Local Similarity	Matches	Conservative	Length
Qy	5.4%	24.9%	100	100	1620;
Qy	4.3	24.9	100	100	1620;
Db	114	114	100	100	1620;
Qy	101	101	100	100	1620;
Db	171	171	100	100	1620;
Qy	159	159	100	100	1620;
Db	233	233	100	100	1620;
Qy	163	163	100	100	1620;
Db	291	291	100	100	1620;
Qy	239	239	100	100	1620;
Db	350	350	100	100	1620;
Qy	378	378	100	100	1620;
Db	400	400	100	100	1620;
Qy	357	357	100	100	1620;
Db	459	459	100	100	1620;
Qy	415	415	100	100	1620;

QQVQVNTGSYEQQHPIGYKLHNKKCQV-----EVTKGL-----

;Accession: A54105
;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
;Molecule type: mRNA
;Residues: 1-2918 <ZHA>
;Cross-references: GB:UD3272
;Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, Y.; Tsipouras, P.
;Nature 352, 330-334, 1991
;Title: Linkage of Marfan syndrome and a phenotype typically related disorder to two differen
;Reference number: S17062; MUID:91304567; PMID:1852206
;Accession: S17063
;Molecule type: mRNA
;Residues: 752-1791
;Cross-references: EMBL:X62209
;Milewitz, D.M.
;Submitted to the EMBL Data Library, December 1992
;Reference number: S31101
;Accession: S31101
;Molecule type: mRNA
;Residues: 752-1407-'R' 1409-1489-1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1926 <IEEE>
;Cross-references: EMBL:X62069
;Genetics:
;Gene: CDB:FBN2
;Cross-references: GDB:128122; OMIM:121050
;Map position: 5q22-5q31
;Superfamily: fibrillin 1: EGF homology
;Keywords: extracellular protein
;Cross-references: 1-29/Domain: signal sequence #status predicted <SIG>
;29/Domain: fibrillin-2 #status predicted <MAT>
;1245-1280/Domain: EGF homology <EGF1>
;1970-203/Domain: EGF homology <EGF2>

Query Match: Score 727; DB 2; Length 2918;
 Best Local Similarity 13.0%; Pred. No. 5.1e-36;

RESULT 4
A57278

Db	1650	I D E C E L P G L C Q G N C I -		N 1657
Qy	622	E R Q E S C G V Q G H - A E - - - - - N O C V S - - - - - C R A G T Y D G A R E C I L C P - - - - - 658		
Db	1668	T F G S F Q C E P Q G Y Y L S E D - T R I C E D - D E C F A H P G V G P G T C Y N T I G N Y T C I C P P E Y M Q N G 1727		
Qy	659	- - - - - N G - T F O N E - - - - -		
Db	1728	G H N C Y D M R K S F C Y R S N G T C E N C E N L P F N V T K R M C C T Y N V G K A N K P C E P T P G T : A D F K 1787		
Qy	652	- - - - - A L K T P E A M N S E C - - - - -		
Db	1789	T I C S N I P G F F D I H T G K A V D I E K E I P G I A N S V C I N Q I G S F R C E C P G F S Y N D L L V C 1847		
Qy	695	- - - - - G 3 - C C - - - - - P G Y S A D S F A P C O L C A L G - ? F C P E A G - - - R T S C F P - - - 73C		
Db	1848	E D I D E C S N G D N L C O R N A D C I N S P O S Y R C E - - - - - C R A G F K L S P N G A C V D R N E C L E I P N 1900		
Qy	731	- C G G G L A T R X G O A T S F Q - D E T R Y Q C S P G H F Y N T T H R C I R - - C P V G T Y Q P E F G N C V S 786		
Db	1901	V C S H G L C V D Q G - S Y Q C I C H N G P K A S Q D Q T M C N D D V D E C B R H P : C G N G T K N T V S Y N C U 1958		
Qy	787	C P G - - - - M T T I D E G S T N I - - T Q C K N B R C G G E L G D F - - - - T G Y I E S P N Y P G N Y P A N T 833		
Db	1959	Y P G F E L T H R N D C S : D E C S S F F G Q V C R N G R C F N E I G S F R C L C N E S Y E L : F G K N C I D T N - 2 C 7		
Qy	834	EC 835		
Db	2018	EC 2019		

Qy	330 RNGGCDH-FCCNTVGSFDGGKKGFULLDEBKSCDVSCLBPTCDH-SCINHPGTFAC 387	Db	1363 DLEDECNSTNSCQHADCKTNGSVC-LKREGYJDFEFTCTG-DECSENLNUCNGQCN 1422
Db	1449 NTNLCENGOC-NVPGAYRCCEMGTTPASDRSCQDIDCSFONICVRSTCNR-PGMFH 1508	Qy	105 IPGNYRCT-CEPGEFLMLDHGHNCLNDECA------LENG 137
Qy	388 ANGRYTL-YGFTHCDDNEES-IN-NGCQQQVCVNTVGSYBQCHPGYKUHWNKDC 442	Db	1423 APGTRCECDDMGPSAAGKACED-DECSPNMIVFGTOHNLPGJFR-ECEIGYELDSG 1482
Db	1509 2CDDGFELDGGNCNPDIEGADPINCNGV-UCNTPRGYBNCNPDDCNATAGVG 1564	Qy	138 G-----CQHTSYNNNGSYECCKEGFFLSSNQHCTH------RS 171
Qy	443 VEVK-GLLPTSVSPVSLHCKSGSDGCFURCHSGIHLSSDTTIRTSVTFKLNEGRKS 501	Db	1483 GNCTDVNECLDPTTISGNVNTPrSSYTCDPFDENLTTRGVCFUTS3GNCYLDIRSFG 1542
Db	1565 VDNVRGNCYLFNGPR-----GDGS-LSCNT----EAGVGVSRSSCC--CS 652	Qy	172 EFG-LSCMN-----KDHGSHI-----187
Qy	502 LKNAELFP-EGURPALPEKHSSVSKESFRYVNLTCSSGKQVPGAGRPSPTPKEMEITYFE 560	Db	1543 DNGDTACSE-GVGVSKASCSCSLCKAKSTGPCE-CPPVNTSEYKILCPEGGEPPNPNTV 602
Db	1603 IGRKAWNPCEPPTVNSTEHTL-----CPGGCFRNP-----ITILLE 1642	Qy	188 -----CKEAPR-----GSVAECECRGPFLANRQDS--I-TCNH---GNGG 223
Qy	561 LETNQEVTASCD-LSCIVTRKTEKLRAIRTLKAHVREQPHQLQSGMNLVOAKKPPRT 619	Db	1603 ILEDIDECQELPGLGCGQSKCINTFGSFQCRCPTGYYLNEDTRVCDDNSETPGICGPg- 1661
Db	1643 DIDEQOELPGICQGNCI-----1660	Qy	224 COHSCDDTADGPECSCHP2QYKMH7GRSCH-REDTVELETESNT-SWGDOKRVKRRLIM 263
Qy	620 SERQABSGVQGH-AE-----NQCVS-----CRAFTYDGARERCILCP-----658	Db	1662 --TCVNTVGNYTCIOPPDYMQVNSGNNNMDMRSLCYDNYYADNQTCTGELLNNTKJM 1778
Db	1661 NTFGSOCECQGQYLSBETRICEDEPAHPGVGPGTCVNTLIGNTYCICPPRYMWN 1720	Qy	284 ETCAVNING-GCDRTCK-----304
Qy	659 -----NG-TFONE-----ALXTPEAMNSEC-----694	Db	1719 CCGSYNIGRAWKPCPQQCP-PISTDDEFTLGSQRPGFVFDYTO-PIVDIDECRIPGVSE 1778
Db	1721 CGHNCMDMRKSFYCYSYNTGTCENBLPNVTKRMCCCTYNKGAKNKP-BPCPPTGTADE 1780	Qy	305 -----HCSCPVGPFQLQDQSKTCKCIDECC-----328
Qy	682 -----ALKTPEAMNSEC-----1781 RTICGNIPGFFDIIHGKAVIDECKEIPGTCANVGCINIGSFRCECPTGFSYNDLJW 1840	Db	1779 NGVCIINMVGSFRCEPVGFFYNDKLKVCEDEDCCNGPVQORNAECAINTAGSYRCDCKEG 1838
Db	695 -----OGLCQ-----PGEFSAQDFAPCQLCAIG-TRFSCFP--RTFSCFP- 730	Qy	329 ---TRNGCDH-----FCRN-YGSFRCGCKGFKLUTDEKSCQDyDECSLD 371
Qy	1841 CEDIDECSCNGNLNLCORNAIDCINSPSYRCE-----CAAGFKLSPNGACVDNECLEBIP 1893	Db	1839 YRFTSTGQCNDRNECEIPNICSQH9QCIDVGSFYCJCHTGFK-NADOTMCID-NCEFB 1898
Qy	731 --CGGLIATRKQGATESEQ-DETRVQCSPGHIFYNTTHCR--CPVGTQYOPERGKONCV 785	Qy	372 RTCDHSCINHPGTFACAHRCYTGFTFGDTCNECSTINGG-CQG-VCVNTVGSYECOC 429
Db	1894 NVCSSHGLCVDLQG-SYQCIQNGFKASQDQTYCMQVDESERHCACTGKNTVGSYNCI 2951	Db	1899 ACGNGTORTGSENTRCHGFIISHNNDCLDVIDCATGNGLNCRGQINTVGSFCGCC 1958
Qy	786 SCPPG-----NTTTDPDGSTM--TOCKNRRCGELGDF----TGYIBSPNYPGNYPAN 832	Qy	43C HFGYKLHWNKKDVCVKGLJ--PTSVSPRVSLHCKSGGDGCFJRCHSIGH:LSSDVITI 487
Db	1952 CYPGPZLTHNDCLDIDPECSFFQVCRNGCFNEIGSKFLCNNEGTEL-PDGKNCITN 2011	Db	1959 NEGYEVAPDGRCTVCDINECLDPRKCA2GT--CONLDGSYRCI--CPRPGYSLONDKE- 2012
Qy	833 TEC 835	Qy	488 RTSVTFKLINE-----GKCSLXNAE-----LFFPLGLR-PALPEKISSKVESPRVN 531
Db	2012 -EC 2013	Db	2023 -----DIEDECVEPEICALGTC--NTEGSFKCLCPDGFSLSSITGRRQDLRMSYCFAK 2064
Qy	RESULTS	Qy	532 LTCSSGRQVPGAPGRBSPTPKEMFIVTEFF-ETNQKEVITASCDS-----CIVYRTER 584
Qy	A5567	Db	2355 F-----EGKKSPPK-----SRNSKQECCLAKGEWMGDPCEL-CPTEP- 21C3
C	Fibrillin I - bovine	Qy	595 LRKAIRLRKAVHRECFFGL-----SGNLDAVKPRTSERQAE-----626
C	Species: Bos primigenius taurus (cattle)	Db	2154 -----CFAFRQ-CPYSSG1LVGPDSAVDNEDECKEDVCHGQJC-NTDGGY 2149
C	Date: 20-Feb-1995 #sequence_revolution 20-Feb-1995 #text_change C2-Aug-2002	Qy	627 --CGVQGQHAEHQCV- SCRAYYYDARERCILCPNGTQNEEGQMTCECP,PR,PGNS 68C
C	Accession: A55567	Db	2150 RCECPFCYLOGNECVM-DECSVGN-----PGN3TCKVIGGEGP-TCEEGFFP 2196
R	Tiistrisa, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byars, P.H.	Qy	681- GAIKTPPEANWNSCGG--LCQGEYKSAQDFAPCQJCAUTSFQPPAGRFS- -PPCGGG 734
R	Genomics 23, 480-485, 1994	Db	2199 GRMYTCB-DINECQNLPLCAFRCVNTSYECK-CPAGYVLE-DRMKCKDECEEG 2254
A	Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to chromosome 12; Reference number: A55567; PMID: 7835920	Qy	735 LATKHOSATSFQOCETRVC-----CSPGHFYNTTHCR1CPV3TYCPBFKGKNCVSCDG 789
A	Accession: A55567	Db	2255 --KHDCAEKOMECKALIGTYLCIGPGWRRPQBEGCV-----DENEQCTKRG 2330
A	Status: preliminary	Qy	793 NTTDDEGSTMITOCKNRQCGGELDFT-----GYIESNPYGNYPANDECTWINP2PK 844
A	Molecule type: mRNA	Db	2351 -----ICNGRCLNTRSSYTCENDGATASPV-----2327
A	Residues: 1-2871 <TIL>	Qy	845 RRIIVVPEIFPLIEDDCGD-----YLMWRKQS5SSNSNTYETCQTYER----888
B	Best Loca.: Similarity 22.7%; Pred. No. 7.6e-34;		
M	Matches 271; Conservative 107; Mismatches 375; Indels 44; Gaps 49;		
Qy	45 PVDECAQGLDDCHADALCQNTPTSYKCSCKPGYQGEGRCQCDIDECGNELUNGGCVHDCZN 104		

Db	2C GEAKRPRTRKTELEHNCAH-----FCINPGSIVCRQGYLLNSDQTTCRQDLCAM 70	Db	718 - PNVCA - -NG----JCSNLLEGSSYMCSCIKHGCTRTDPDHKHCRDIDECCQSNLCVNGQCN 769
Qy	93 ELNGGCVHDCLNIPGNYRCTFDGFMLAHDGHNCNLVDVECLENNGGCOHTCVNNGSYEC 152	Cy	341 IVGSFDGCKKGFKLUTDEKSQDVDEBGSUDRTCH-SCTNHPGTACAGNRGTTJYGF- 398
Db	71 E-DHNEQLCVNPGSIVCQTSYGAELAEDEKRCAVADYGASENHGCEBEVNADGSTC 129	Db	770 TEGSFRTCGCYQLSAAKDCDCE-DECQHRLCAGQCANTEGSFOCQCDGVRASGLG 829
Qy	153 CCKEGFPLSDNQHTCHRSSEGSLSMKNDGICSHCKEARGSVACECRPFELAKNQD 212	Cy	399 THCGDTNECSNTNGGQQ-VCNTVGSYECOCHPSYKLHNK-----KCDEVVKGJLPTS 452
Db	130 QCGEFGALNPDKTCKMD----YCASSNMCCQHECVNTD-DYSCHCLKGFTLNPDXT 184	Db	830 DAGEDNECLBDPKSV-CRGDINTAGSYDC-CPDFQJDNKRTCODNB-EHPOL--- 884
Qy	213 C-ILTCNHNGGCGCHSDDADGPECSCHQYKMTDGRSCLEREDTVLVEVTEANTS 270	Qy	453 VSPRVLSHGKSG--GGDGF-LRCHSGHLSSDVTITSTVTFKLNEGKCSLNAELF 508
Db	185 CRRINYCALNKPGCENHECVNNEESYCRCHRGYTLDPNGKTCR-----228	Db	885 -----CGPGECLNT-EGSFCHVQQGFISADQRT-----CEDI 918
Qy	271 VDGDKRVKRLIMETCAVNNGCDRTCKDTSGVHSCEPVGFTLQDGKTKDIDECCTR 330	Qy	509 PEGLRPAPPERH---SSKEVESTRYNLTCSGKGQYB-----GAGRGP- 547
Db	229 -----VDHQAQQDGCQEQLCLNTDSFVQCSEFL-NEDLKTCRSRYCYCLS 276	Db	919 DECVNNTVCSHGFCDNTAGSFRCI--CYQFCAPQDGQCVDNECELLSGVGEAFC 975
Qy	331 NGCDFHCKRNIVGSFDCGCKGFKLUTDEKSQDVDECSU-DRICDCHSICNHPGTFACAC 389	Cy	548 STPKEMITV-----EFELTNQEVTAASCDLSCIVKVR-ERKURKAITLURKA-----595
Db	277 DHGECSCVNNDRSFAQCPGQPHVLSRGKTCALKDSDGKGEHSCVSSEDFVQOC 336	Db	976 ENVEGSEFCAADENCBYSPTVQCSRSRSTSDUDCQDQVQPEEKKCYVNLNDASJZCNV 1035
Qy	390 NRGTYL-GFTHGDTNECSINNGCQVQCPYKJHANKRDKCVCENKG 447	Qy	536 -----VREQFHQLS--GXNLDVAKKPRTSERQAEQSGVQGCH-----AEN- 636
Db	337 PEGYLREDGKT-CRKDVKQDAIDGCEHICVNSDDSYTCECLEGFLRAEDGIREC-RRKD 394	Db	-036 LAPNVTQECCTSGAGNCONEIFCPVPLGTAETMPCKGKFVPAGESSEAGGENV 1095
Qy	448 ILPTSVSPRVSLHCGKTS---GGGDGCFVLRHSGZHLSSDVTIRISTVFKLNEGKCS- 502	Qy	637 -----SCRAYHVDGAEERC-----LCPNGF 662
Db	395 VCK-----STHAGCEHIVNNGSYICKCSBEGFVLAEDGRRCK---KCTEGPJD 441	Db	.096 KDAECLLFGQEICKNGFCLNTRPGYECYCKQGTYDPPVXLQCFDMCQDPSSCIDGQC 1155
Qy	663 ONEEGQNC-----EPCEPGRNS-----3 681	Qy	663 -----3 681
Db	1156 VNTEGSINCFCIHPMMI.DASPERKRCRPAESREQIE.DVYQJCLMREHSDYEVCSRPLV 1215	Db	1156 VNTEGSINCFCIHPMMI.DASPERKRCRPAESREQIE.DVYQJCLMREHSDYEVCSRPLV 1215
Qy	682 AJK-----PEANVMSECQGLCPGEYESADGPAQCOLCALGFPQERGRTSFC-----729	Qy	682 AJK-----PEANVMSECQGLCPGEYESADGPAQCOLCALGFPQERGRTSFC-----729
Db	1216 KQT-TYTCBCCDLYGEAQM-QC-ALC-PLKDDDDYAQCNIPVTRGRQ-YGRDALVDFSE 1271	Db	1216 KQT-TYTCBCCDLYGEAQM-QC-ALC-PLKDDDDYAQCNIPVTRGRQ-YGRDALVDFSE 1271
Qy	736 -----PCGGSLATKHOQATSPOGCTVQCSPOPHVNNTTHICLFCVG-----TYQPF 779	Qy	736 -----PCGGSLATKHOQATSPOGCTVQCSPOPHVNNTTHICLFCVG-----TYQPF 779
Db	1272 QYTPREADPYF-QDRELNSFEELQAE-EC-3LNGENGSRVQEGYTCQCDJQYHLCT 1328	Db	1272 QYTPREADPYF-QDRELNSFEELQAE-EC-3LNGENGSRVQEGYTCQCDJQYHLCT 1328
Qy	790 GKNCNVSPGNTTDDFGSTN-TCQFNRRCGEGLGF-----TGYIES-PNY-PNY 8229	Qy	790 GKNCNVSPGNTTDDFGSTN-TCQFNRRCGEGLGF-----TGYIES-PNY-PNY 8229
Db	.229 ARNTCFD---VNECDELNRKSLCKRANKINTDSYKLCJPOVPSDKPNYCTPLK- 1383	Db	.229 ARNTCFD---VNECDELNRKSLCKRANKINTDSYKLCJPOVPSDKPNYCTPLK- 1383
Qy	830 PANTE 834	Qy	830 PANTE 834
Db	1384 AENLE 2395	Db	1384 AENLE 2395

RESULT 9

A35626 transforming growth factor beta-1-binding protein - human

C;Species: Homo sapiens (man)

C;Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 11-Jan-2000

C;Accession: A35626

R;Kanazaki, T.; Olofsson, A.; Moren, A.; Wernstedt, C.; Heijman, U.; Miyazono, K.; Claess

Cell, 61, 1051-1061, 1990

A;Title: TGF-beta1 binding protein: a component of the large latent complex of TGF-beta1

A;Reference-number: A35626; MUID:9275601; PMID:23507033

A;Accession: A35626

A>Status: Preliminary; not compared with conceptual translation.

A;Molecule type: mRNA

A;Cross-references: GB:34057; NID:9339547; PID:9339548

C;Superfamily: unassigned EGF-related proteins; EGF homology

C;Keywords: alternative splicing <EGF>

P;750-791/Domain: EGF homology <EGF>

Query Match 11.1%; Score 624.5%; DB 2; Length 1394;

Best Local Similarity 24.9%; Pred. No. 3.9e-30;

Matches 240; Conservative 94; Mismatches 334; Indels 297; Gaps 48;

C;Masking Protein precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Accession: A32261; Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 21-Jan-2000

R;Tsujii, T.; Okada, F.; Yamaguchi, K.; Nakamura, T.

Proc. Natl. Acad. Sci. U.S.A. 87, 8835-8839, 1990

A;Title: Molecular cloning of the large subunit of transforming growth factor type bet

A;Reference number: A32261; MUID:91062373; PMID:2247454

A>Status: Preliminary

A;Molecule type: mRNA

A;Residues: 1-1712 <TSU>

A;Cross-references: GB:MS5431; NID:9207285; PID:AAA42235; 1; PMID:9207286

C;Superfamily: unassigned EGF-related proteins; EGF homology; EDL receptor ligand-bind

P;911-947/Domain: EGF homology <EGF>

Query Match 10.6%; Score 592; DB 2; Length 1712;

Best Local Similarity 24.3%; Pred. No. 4.5e-28;

Matches 230; Conservative 102; Mismatches 305; Indels 31C; Gaps 52;

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein in rat							
45 DVDECAQGLDDCHADALCONTPSYKCSCKPGY-GEGEPROCEDIDE-----GNEL 94	Y						
907 DIDECAQAHLC-SQRGRCNTEGSFLCIPAGFAESGSNC-DVDEELRPPVCRDRCI 965	c						
95 N-----GGC-----VHDCLNTPGNYRCT-CFDGFMLAH 221	Y						
966 NTAGAFRCEYCDSSYRMSRGHEDIDECLTPSTCPPEQCNPSGYQCVPCTEGFR--- 1022	c						
122 DGHN-----CLDVDBCLEANNGCQHTCCTVNGSYECCKEGFFLNSNCHTC--IIRSEGLSC 177	Y						
1023 -GRNGQCLDVDECLQPKVCTNGSCNLBSSYMSCHKGYSPPEDHRHQDIDECQQGNLC 1081	c						
178 MNKDHGCSHICKEAPRGSYACECRGPFEU-AKNDQDILTENHNGGCOH-SCDIDTDG 234	Y						
1082 MNQG-----CXTND-GSFRCCTCGQYQISAAQDCEDIECHRH-LCSHQCRNTEGS 1133	b						
235 PECSCHPOYKHMTHGDSCLEREDTVELETSNTSVVDGDKVKRLLMETCAVNNGCD 294	Y						
1134 FQCLNQGYRASVLGDHC---ED---INE-----CLEDSSVCQCG-- 1167	c						
295 RTCKDTGSTGVHCSCPVGFTLQLDGKTKDIDECOTRNNGCDHE--CKNIVGSSFDGCKKG 352	Y						
1168 -DCINTAGSYDCTCPDGQLN-ENKGCDPINEC-AQPGICAPHGECCLNTQGSHFCVCEQ 1224	c						
353 FKLLTDEKSDQVDECSLDRTCD-HS-CINHPGTACACNRGTYL-YGFTFHGDTNECSI 409	Y						
1225 FSISADGRTCEDIDECAVNTVCDSHGFCDNNTASFRCLCYQGFQAPDQGCVUDNECEL 1284	c						
410 NNGGGC-QQCVNNTVGSYEQCHPGYKLH-----WNRKDC---V 443	Y						
1285 LSGVGEACANTEGSFLCVADENQEYSPMTGCSRATEDGSDVDROPKEERKEYCYNL 1344	c						
444 EVKGLPLPTSVSPRVS LH-CGKSGGGDGFCLRCHSGIHSUSSDVT-TIRTSTFKLNBKCS 50:	Y						
1345 NDASLCDNVLAAPNTYKOECCCTSGAGNW----- 1372	c						
502 LKNAELFP-----EGLRPAPEKHSSYKESFR-----YVNLTCSGK 538	Y						
1373 -DNECIEPPCPVQGTAEPSEMCPRGKGFPAESEYYCETGEGY 1431	c						
539 QVPAGPGRSPSTPKMFIYFELETNOKEVTASDLSCIVKRTERLRKAIRTLRKAVHR 598	Y						
1432 CLNTQPGYECYKE-----GTYDEVKLQ----- 2456	c						
599 EQFHQLQSMNLDYAKKPPRTSERQAESCGVGQCHAENQCVS-----CRAGYXDG 650	Y						
1457 -----FDMDEQDQP-----NSC-----IDGQCVNTDEGSYNCFCNHRMVIDAS 1493	c						
651 RERCILCPNPTFNEQNTCEP-----PREPNSGALKT-----PEANXSE 693	Y						
1494 EKRCVQPTSSNEQIETDVGDLQWEHLSEEYYCSPRLVKQTYYBCCLYGEANGM-Q 1552	b						
694 CGGLQPGYSEADGFAPCOLCALGTFOPEAEGRTSCF-----PCGGSLATKQ 740	Y						
1553 C-ALCPMKD-----SDYA---OLCNI----FVTGRRPYPGRDALVDFSEOYGPETPYFQDR 1603	c						
741 GATSFQDCETRQVQSPGHFYNTTHRCLRCRCPVG-----TYQPEFKANCYCSPGNTT 793	Y						
1604 FLSNSEEEQAE-EC-GILNGCENGRCVRQEQSYTCDFDG3YHLDMAKMTCV-CV-NECS 1658	c						
794 DPDGSTNTQCKNRGGE-GDF-----TGYIES-----PNYPGNYPANT 833	Y						
1659 ELNNRMSL-CKNRACKINTEGSYKCVCLPGYVFSEKPNY-CTPLNT 1761	b						
RESULTS: 12							
A55444							
1 latent transforming growth factor-beta-binding protein - human							
C:Species: Homo sapiens (man)							
C;Date: C3-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 11-Jan-2000							
R:Accession: A55444							
R:Moren, A.; Olofsson, A.; Sternman, G.; Sahlin, P.; Kanazaki, T.; Claesson-Welsch, U.; tr. J. Biol. Chem. 269, 32469-32478, 1994							
A:Title: Identification and characterization of LTBP-2, a novel latent transforming growth factor-beta-binding protein							
A:Accession: A49457							
A:Title: Sequence revision 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Sep-2002							
A:Accession: A49457							
A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein in rat							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage by matrix metalloproteinase-2							
A:Accession: S74095							
A:Title: Structure and expression of fibulin-2 tc cleavage							

A;Accession: A55494	Qy 779 PGKNNCYSCPGNTTDFGSTMITQCKNRGGELGDFGYIESNYPGNYPANTEC	835
A;Status: preliminary		
A;Molecule type: mRNA		
A;Residues: 1-1820 <MOR>		
A;Cross-references: GB:Z77976		
C;Genetics:		
A;Gene: GDB:LTPB2		
A;Cross-references: GDB:568901		
A;Map position: 11pter-11qter		
C;Surfprotein: unassigned EGF related proteins; EGF homology		
F:122-1257/Domain: EGF homology <EGF>		
F:1525-1565/Domain: EGF homology <EGF1>		
Query Match Score 10.24; Best Local Similarity 23.2%; Matches 208; Conservative 110; Mismatches 339; Indels 240; Gaps 45;		
Query 46 VDECAAGLGDODCHADALCQNTPTSYKCSCKPQYGEGRQ-CEDIECCGNELNGGCVHD- 101		
Db 845 ZDRCAAGATAVNCGPCTCVNLPGTFCVCSPYQLHPSQAYCTDNEC--LRDPQCQGKR 90-		
Qy 102 CLNIPGNYRCYCFDGFMLAHDG--HNCLDVDECLENNGGQ-HTCVNMGSYECCKCSGP 158		
Db 902 CINRVSYSSCYPGTYLATSGATOEQCDINEC-BQPGVCGGGQCTNTBGSYHECDCQY 963		
Qy 159 PL-----SDNQHTCI----HRSEEGLSCKMRD-----L 161		
Db 961 IMVRKGKHCQDINECRPGTCPDGCVNSPQGSGTCLACEESYRGSG-SCVDVNECLTPGV 1019		
Qy 182 -HG-GSHICKEAPRGSVACRGPFELAKNQRDCILTNHNGKQCOHNSCDTAAQGPEC- 237		
Db 1020 CAEGKCTNL---EGSFRSGCEQGVETSDEK-----GCQ-DVDECASRAS.P 1062		
Qy 239 -----SCHPOYKMHMTDGRSLCLEDRTYLEV-TESN-TSWVDDGDKRTRKRLM 283		
Db 1063 TGLCUNTEGSSFACSAENG/TWNEDTAC--ED-----L 1094		
Qy 284 ETCAVNRGGCORTCKDTSTGTHC-SCPVGFTLQLDGKTCKDIDEQTRN3GC-DHFCRN 34;		
Db 1095 DECAFPGVCPSPGVCTINTAGSSPKQCDGGYPRSPLGSDCVDDEPOSSCLGEGCINT 1154		
Qy 342 VGSFDGCKKGFKPLLTDEKSQDVDECSLDRTC--DHSCINHPEQFACAGNRGY-TLYGF 398		
Db 1155 VGSYQCLCPQFQ-LANGTVEDVNEMGEHHCAPHECLNSHGSFFCLCPGFVSAEGG 1213		
Qy 399 TRCGDTNECS-NNGGQQCVNTVGSYECOCHGKLYHNNKKDCWEVKGSLPLPTSVPRTS 456		
Db 1214 TSCQDVDECATDPCVGHCNTEGFNCLGTGPSPSGEVCYDIDECEDYGDPPVCGT 1273		
Qy 459 LHCKGSGGGDCFCFLRCHSGTHLSSDVTITTSVTKLNEGCKSLKNAELFPPEGLRPALPE 518		
Db 1274 WRCENSPGSPSYCRLGQPGFMAPNGPDCI-DIDECANDTMG----- 1314		
Qy 519 KH--SSVKEFSRYVNLTCSSKGKVQVGPAPGFPSTPKEMFTVEPEL-----ETNOKE 567		
Db 1315 SHGFCNDTGSRCL---CDQG-----FEISPQGMCDVNECE 1350		
Qy 568 YTASCDLSCIYKTRKLKAIRTRKAVHREQFHQLSGNLYQAKPFR--TSEHQAE 625		
Db 1351 LMLAVCGAALCENVEGSFLCLASDLEYDQEGTAHGGLEVCRLPORDHAPAPRM 1410		
Qy 626 SCVGQ-GHA-----ENOCVSCRAGTYDGAERERCILCPNGTFQNNEGOMTCP 673		
Db 1411 DCYSGQRGHAPSSVLGRNTIAECCCTQATWGDA---CDLQCLP-----SEDAEFSII 1461		
Qy 674 CRRPGNSGALKTPPEAW----NMSEC----NMSEC----GGLQGPGE--YSDGFAPCOLCALGTF 719		
Db 1462 C-PS3KGYIPIVEGANTFGQMTYTDADECVLFPGPGLCPNRCRCLNTVPGYV-C-LCNPQ-F 1516		
Qy 722 QPBAGRTSCFPGGGLATKHSQGATSFQDCETRVOCSPGHEYNT-THRCCRCPVGYTQPE 778		
Db 1517 HYDASHKKC-----EDHDECQD-JACENGECVNTEGSFHFCSPFLTJ--D 2559		
Qy 29 779 PGKNNCYSCPGNTTDFGSTMITQCKNRGGELGDFGYIESNYPGNYPANTEC	835	
Db 1560 LSQQCRCKMNSTSSETED-LFDHDHMDCWKK-----VNNDVCSEPLRGRRTYTEC	605	
RESULT 13		
S80404C fibulin, splice form C precursor - mouse		
N;Alternate names: basement-membrane protein BM-9C		
C;Species: Mus musculus (house mouse)		
C;Date: 24-Jul-1998 #sequence revision 24-JUL-1998 #text_change 02-Aug-2002		
C;Accession: S70404; S78560; S36450		
R;Pan, T.C.; Klug, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.L.		
Eur. J. Biochem. 215, 733-740, 1993		
A;Title: Sequence of extracellular mouse protein BM-9C/fibulin and its calcium-dependence		
A;Reference number: S34968; XJID:9338897; PMID:8354280		
A;Accession: S70404		
A;Molecule type: mRNA		
A;Cross-references: EMBL:X70854		
A;Residues: 1-685 <PAN>		
A;Crosses: EMBL:X70854		
C;Genetics:		
A;Introns: 568/3		
C;Superfamily: fibulin-1; EGF homology		
C;Keywords: alternative splicing; basement membrane; calcium binding; extracellular ma		
F:1-29/Domain: signal sequence #status predicted <SIG>		
F:30-629/Product: fibulin, splice form C #status predicted <MAN>		
F:98, 537, 541/Binding site: carbohydrate (Asn); (covalent) #status predicted		
A;Molecule type: mRNA		
A;Residues: 1-39; 'P' 41-685 <CHU>		
A;Crosses: EMBL:X70854		
Query Match Score 10.24; Best Local Similarity 23.2%; Matches 208; Conservative 110; Mismatches 339; Indels 240; Gaps 45;		
Query 46 VDECAAGLGDODCHADALCQNTPTSYKCSCKPQYGEGRQ-CEDIECCGNELNGGCVHD- 101		
Db 845 ZDRCAAGATAVNCGPCTCVNLPGTFCVCSPYQLHPSQAYCTDNEC--LRDPQCQGKR 90-		
Qy 102 CLNIPGNYRCYCFDGFMLAHDG--HNCLDVDECLENNGGQ-HTCVNMGSYECCKCSGP 158		
Db 902 CINRVSYSSCYPGTYLATSGATOEQCDINEC-BQPGVCGGGQCTNTBGSYHECDCQY 963		
Qy 159 PL-----SDNQHTCI----HRSEEGLSCKMRD-----L 161		
Db 961 IMVRKGKHCQDINECRPGTCPDGCVNSPQGSGTCLACEESYRGSG-SCVDVNECLTPGV 1019		
Qy 182 -HG-GSHICKEAPRGSVACRGPFELAKNQRDCILTNHNGKQCOHNSCDTAAQGPEC- 237		
Db 1020 CAEGKCTNL---EGSFRSGCEQGVETSDEK-----GCQ-DVDECASRAS.P 1062		
Qy 239 -----SCHPOYKMHMTDGRSLCLEDRTYLEV-TESN-TSWVDDGDKRTRKRLM 283		
Db 1063 TGLCUNTEGSSFACSAENG/TWNEDTAC--ED-----L 1094		
Qy 284 ETCAVNRGGCORTCKDTSTGTHC-SCPVGFTLQLDGKTCKDIDEQTRN3GC-DHFCRN 34;		
Db 1095 DECAFPGVCPSPGVCTINTAGSSPKQCDGGYPRSPLGSDCVDDEPOSSCLGEGCINT 1154		
Qy 342 VGSFDGCKKGFKPLLTDEKSQDVDECSLDRTC--DHSCINHPEQFACAGNRGY-TLYGF 398		
Db 1155 VGSYQCLCPQFQ-LANGTVEDVNEMGEHHCAPHECLNSHGSFFCLCPGFVSAEGG 1213		
Qy 399 TRCGDTNECS-NNGGQQCVNTVGSYECOCHGKLYHNNKKDCWEVKGSLPLPTSVPRTS 456		
Db 1214 TSCQDVDECATDPCVGHCNTEGFNCLGTGPSPSGEVCYDIDECEDYGDPPVCGT 1273		
Qy 459 LHCKGSGGGDCFCFLRCHSGTHLSSDVTITTSVTKLNEMEGCKSLKNAELFPPEGLRPALPE 518		
Db 1274 WRCENSPGSPSYCRLGQPGFMAPNGPDCI-DIDECANDTMG----- 1314		
Qy 519 KH--SSVKEFSRYVNLTCSSKGKVQVGPAPGFPSTPKEMFTVEPEL-----ETNOKE 567		
Db 1315 SHGFCNDTGSRCL---CDQG-----FEISPQGMCDVNECE 1350		
Qy 568 YTASCDLSCIYKTRKLKAIRTRKAVHREQFHQLSGNLYQAKPFR--TSEHQAE 625		
Db 1351 LMLAVCGAALCENVEGSFLCLASDLEYDQEGTAHGGLEVCRLPORDHAPAPRM 1410		
Qy 626 SCVGQ-GHA-----ENOCVSCRAGTYDGAERERCILCPNGTFQNNEGOMTCP 673		
Db 1411 DCYSGQRGHAPSSVLGRNTIAECCCTQATWGDA---CDLQCLP-----SEDAEFSII 1461		
Qy 674 CRRPGNSGALKTPPEAW----NMSEC----NMSEC----GGLQGPGE--YSDGFAPCOLCALGTF 719		
Db 1462 C-PS3KGYIPIVEGANTFGQMTYTDADECVLFPGPGLCPNRCRCLNTVPGYV-C-LCNPQ-F 1516		
Qy 722 QPBAGRTSCFPGGGLATKHSQGATSFQDCETRVOCSPGHEYNT-THRCCRCPVGYTQPE 778		
Db 1517 HYDASHKKC-----EDHDECQD-JACENGECVNTEGSFHFCSPFLTJ--D 2559		
Qy 29 779 PGKNNCYSCPGNTTDFGSTMITQCKNRGGELGDFGYIESNYPGNYPANTEC	835	
Db 1560 LSQQCRCKMNSTSSETED-LFDHDHMDCWKK-----VNNDVCSEPLRGRRTYTEC	605	

Db	615 PG-----DSQIATAGNEFFTRKV-----SNSGV----	643	Db	436 GRSCEDINECS--SSPCSGECANYGSYCCRCYCLSDVDGVTCECDECA;PTGGHI 493
Qy	528 RYVNLTCSQRQVPSAPGRSTPKEMIFTEFEL 561		Qy	374 CDHSCINHPGTACAC-NRGTYL: GPTHGDTNEC--SINK3QQVQVNTGSYEC-- 427
Db	644 -VALT---KPIP-----EPDJLLIVKMDL 664		Db	494 CSYRCINIPSSFOCSPSSGYRLAPNGRNCQIDECVTGHNCs-NETCFN1QAFRCA 553
	RESULT 14		Qy	428 -QCHPGYK-----LHNWKDCVEVKLL----PTSV-SPVSLHGKSGG 467
C3646	fibulin 1 precursor, splice form C - human		D2	554 FECPENYRSGATRERLPHEN-RECSKPLR-TYH:SFPTN:QAPAVFRXGPSSAV 612
N;Alternate names: fibulin C			Qy	468 DGCFLRCHSGHLSSDVTITSTFKLNKEGKCSLKNAL:FPGLRPLPEKHSSVKEF 527
N;Contains: fibulin 1 splice form A; fibulin 1 splice form C			Db	613 PG-----DSNQLAITGNGEEFFT:RKV-----SPHSGV---- 641
C;Species: Homo sapiens (man)			Qy	526 RYNNLICSSKGKVPGAPGRGSTPKEMIFTYEFEL 561
C;Date: 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 02-Aug-2002			Db	642 --VALT---KAVP-----EPFDLLTVYKMDL 662
R;Accession: GB:X53741; A36346; A36246; A32B36				
R;Graves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.				
J;Cell Biol 111: 3155-3164, 1990.				
A;Title: Fibulin is an extracellular matrix and plasma glycoprotein with repeated domain				
A;Reference number: A36346; MUID:91100426; PMID:2263669				
A;Accession: C36346				
A;Molecule type: mRNA				
A;Residues: 1-683 <ARG>				
A;Cross-references: GB:X53743; NID:931416; PID:CAA3772.1; PID:g31419				
A;Accession: A36346				
A;Molecule type: mRNA				
A;Residues: 1-566 <AR2>				
A;Cross-references: GB:X53741; NID:931414; PID:CAA3770.1; PID:g3-415				
A;Gene: GDB:FB:BN1; FBLN1				
A;Protein Family: fibulin-1; EGF homology				
A;Keywords: alternative splicing; glycoprotein				
F:30/63/Domain: signal sequence #status predicted <SIG>				
F:30/63/Domain: fibulin 1 splice form C #status predicted <MBT>				
F:186-214/Domain: EGF homology <EG>				
F:98-535/539/Binding site: carbohydrate <EGII>				
Query Match 10.1% Score 565; DB 2; Length 683;				
Best Loca: Similarity 27.9%; Pred. No. 7.6e-27; Mismatches 77; Indels 192; Gaps 31;				
Matches 160; Conservative 77; Mismatches 155; Indels 192; Gaps 31;				
44 EDVDEAQGLDCHADALCONTPSYKC---SCKPGYQ_GEGRGCDIDECHGELNSGC 98				
215 EDVNECITGSSCRGESCINTVGSRCDISSCTGYELTEJNSCKD-DECSEGII-NC 273				
99 VHD--CLNIPSNYRC---TQFDGEMLAHQHNCIDINECSISAPCPIGHTCNTESYY 150				
Db	274 IPDFIFICNTLSFRCPKQLQKSGFQDLAG-NCIDINECSISAPCPIGHTCNTESYY 332		Qy	44 EDVDEAQGLDCHADALCONTPSYKC---SCKPGYQ_GEGRGCDIDEBCG---NELNG 96
Qy	151 EC----CCRGFFPSDNQHTC1HNRSEBGSCKMNDH-----GCHICKSAPRGSYA 197		D2	718 EDNECYTDLHTCSRGEHNTLGSFHCHYKALTCPGALKDGESEDVDECAMGTH:CCP 777
Db	333 TCQKNPVNCGRY-----HUNNEGTRCVYDCECAPAEPGKGHRCVNSP-GSFR 393		Qy	97 SCVHDQNLNPVNYRC---TCFDGMLAHDGNCNCDVBDOLENGGCQ -H7CUNVMGSY 150
Qy	198 OCECPGPFELANQRDCLLTCNPNHGNGCOHSSCDTAA:GESSCHPOYKHTDGRSCLERED 257		D2	778 GEJ---CQTKGSFYCOARQRMGDFQDPEG-NCDVNBETSLSEPCRGFGSCINTVGSY 334
Db	382 CECKTGYYFDGISRMVY-----DVNEQRYP-----GRIC----- 411		Qy	151 EC----CCKEGFFLUSDNCHTC1HRSSEGLSCLMRKDGGC---SH7KEARPGSVACECRPG 203
Qy	258 TVLEVTESTSNTSVWDGDKRVKRLLMETCAWNNGCORTSTGVHSCPYGFTLQD 317		Db	835 TQJRNPLIACRGYHSDDGAKCVDNE---CEGVHRRGEGQVCHNLP-GSYRDKAG 389
Db	412 -----CHK-----SENLTGSYLSCSVCGRFSYD 435		Qy	264 FJLAKNORDCILCTCNHGNGCQHQSCTADGPECSCHPQYKMH:TGESSLEREDTVLETF 263
Qy	318 GKTCKDIDECCQTRNGGCDHFKNIVGSPDCCCKGFKL-TDEKSCQDVECSL---DRT 373		Db	890 F-----QRZD----- 901

Qy 264 ESNNTSVVDDKDKRVRKRRILLMETCAVNNGG-CDRTCKDTSTGVHCSGPVGFTLQDGKTKC 322
 Db 902 -----VNECWASPRGLRCORTCENTLGSYRCASGFLLAAGDKRCE 942
 Qy 323 D-DECOTRNGGDHFCKNIVGSEDCGKKGFKLLTDEKSCDPODECSLDR - TCHSCIN 380
 Db 943 DYNCEAQQR--CSQECANIYGSYOCYCRQVOLAEDGHTCDIDEAQAGAILCFRCLN 1000
 Qy 381 HPGTFACAC-NRGYTL YGFTHGDTNECSINNGGCO --QVCVNTVGSYEC --OCHPGY 433
 Db 1001 VPGSYQZACPESGYMTANGRKDVDECATGTHNSBAETCHNTJGSFROLRFCCPNY 1360
 Qy 434 KJHWNKKDCVVEVKGLIPLTSVRSVLHGKSGGGCFLRHSGLSSDVTTTTSVTF 493
 Db 1061 -VQVSPTPKC-----ERTTCH-----DFLZCQNS ---PART-----HY 1089
 Qy 494 XNEGKCSLKNAELFPGLRPA 515
 Db 1090 QLNFTQGLVPAHIFRIGPAPA 1111

Search completed: October 22, 2003, 11:31:11.
 Job time : 56 secs

Scoring table:	BLOSUM62	Gapop 10.C , Gapext 0.5				
Searched:	127863 seqs,	47026705 residues				
Total number of hits satisfying chosen parameters:	127863					
Minimum DB seq length: C						
Maximum DB seq length: 2000000000						
Post-processing: Minimum Match C ³						
Maximum Match 100%						
Listing first 45 summaries						
Database :	SwissProt_41::*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
Result No.	Score	Query Match Length DB ID	Description			
1	727	13.0 2911 1 FBN2_HUMAN	P35556 homo sapien	RN	SEQUENCE OF 752-15CS FROM N.A.	
2	724	12.9 2907 1 FBN2_MOUSE	Q61555 mus musculus	RX	MEDLINE:91304567; PubMed:552206;	
3	692	12.3 956 1 MTN2_MOUSE	Q08746 mus musculus	RA	Zhang H., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M., Tsiprassos P., Ramirez F., Hollister D.;	
4	691	12.3 2871 1 FEN1_BOVIN	P98113 bos taurus	RA	"Linkage of Marfan syndrome and a phenotypically related disorder to FBN2."	
5	691	12.3 2871 1 FEN1_HUMAN	P36555 homo sapien	RA	two different fibrillin genes.";	
6	684.5	12.2 956 1 MTN2_MOUSE	Q00539 homo sapien	RN	Nature 352:330-334 (1991).	
7	684.5	12.1 2871 1 FEN1_PIG	P22064 homo sapien	RX	VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.	
8	624.5	11.1 1394 1 LTBS_HUMAN	Q14766 homo sapien	RA	MEDLINE:96003599; PubMed:793032;	
9	624.5	11.1 1395 1 LTBL_HUMAN	Q8c918 mus musculus	RA	Purram B.A., Zhang H., Ramirez P., Milewicz D.M.;	
10	624.5	11.1 1389 1 LTBS_MOUSE	Q8c919 mus musculus	RC	fibrillin-2 (FB2) mutations result in the Marfan-like disorder, congenital contractual arachnodactyly.";	
11	601	10.7 1713 1 LTBL_MOUSE	Q00918 rattus norvegicus	RL	Nat. Genet. 11:456-458 (1995).	
12	601	10.7 1712 1 LTBL RAT	P37889 mus musculus	RN	VARIANTS CCA HIS-1114, MEDLINE:98407739; PubMed:9737771;	
13	592	10.6 1221 1 FBBL_MOUSE	P23142 homo sapien	RA	Babcock D., Gasner C., Fracke U., Maslen C.,	
14	577.5	10.3 1221 1 FBBL_RAT	Q73775 gallus gallus	RA	"A single mutation that results in an asp-to-his substitution and partial exon skipping in a family with congenital contractural arachnodactyly."	
15	567	10.1 703 1 FBBL_HUMAN	Q9r172 rattus norvegicus	RT	Hum. Genet. 103:22-28 (1998).	
16	562	10.0 704 1 FBBL_CHICK	P08055 homo sapien	RT	VARIANTS CCA PHE-1141 AND TRP-1252.	
17	560	10.0 184 1 FBBL_HUMAN	Q08819 mus musculus	RA	MEDLINE:20259236; PubMed:0797116;	
18	558	10.0 705 1 FBBL_MOUSE	P07207 drosophila melanogaster	RA	"A single mutation that results in an asp-to-his substitution and partial exon skipping in a family with congenital contractural arachnodactyly."	
19	529.5	9.4 798 1 FBBL_CAEEL	Q77469 xenopus laevis	RT	"Two novel fibrillin-2 mutations in congenital contractural arachnodactyly."	
20	526.5	9.4 2321 1 NTC3_HUMAN	Q9ummt7 homo sapien	RT	Am. J. Med. Genet. 92:7-12 (2000).	
21	524.5	9.4 598 1 FBBL_CERAE	Q8mij9 cercopithecus aethiops	RT	FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
22	524.5	9.4 2318 1 NTC3_MOUSE	Q61982 mus musculus	RT	THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
23	516.5	9.2 2319 1 NTC3_RAT	P31655 mus musculus	RL	CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL	
24	509.5	9.1 2703 1 NOTC_DROME	Q07638 rattus norvegicus	CC	Am. J. Med. Genet. 92:7-12 (2000).	
25	501.5	8.9 2524 1 NOTC_XENLA	Q04212 brachyrhynchus	CC	FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
26	500.5	8.9 2470 1 NTCC_MOUSE	Q35516 mus musculus	CC	THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
27	500	8.9 2470 1 NTCC_BRARE	P46531 homo sapien	CC	CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL	
28	489.5	8.7 1964 1 NTCC_MOUSE	P46530 brachyrhynchus	CC	Am. J. Med. Genet. 92:7-12 (2000).	
29	489	8.7 2531 1 NTCL_RAT	Q07638 rattus norvegicus	CC	FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
30	489	8.7 2471 1 NTCC_HUMAN	Q04211 homo sapien	CC	THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
31	485	8.7 2556 1 NTCL_HUMAN	P46531 homo sapien	CC	THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.	
32	484.5	8.6 2437 1 NTCL_BRARE	P46530 brachyrhynchus	CC	CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL	

CC ARACHNODACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARfan SYNDROME, BUT DOES NOT EFFECT THE
ACRA AND THE EYES.

CC !- SIMILARITY: Contains 47 EGF-like domains.

CC !- SIMILARITY: Contains 7 TGF-Beta binding protein (TGFBBP) domains.

CC !- DATABASE: NAME=Elasic Fiber Homepage; NOTE=Fibillina 2 page;
WWW=<http://ee.wustl.edu/Genes/FBN2.htm>.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
use by non-profit institutions as long as its content is in no
modified and this statement is not removed. Usage by and for commercial/
entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U03272; AAAI8950.1; -.
DR EMBL: X62009; -; NOT_ANNOTATED_CDS.
DR HSSP: P35555; 1EMN.

DR Genew: HGNC:3604; FBN2.

DR MIM: 121050; -.
DR GC: GO:0005578; C:extracellular matrix; TAS.
DR GO: GO:0005201; Extracellular matrix structural constituent; TAS.
DR GO: GO:0007345; Embryogenesis and morphogenesis; TAS.
DR InterPro: IPR000152; Asx hydroxyl.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001438; EGF II.
DR InterPro: IPR0062209; EGF like.
DR InterPro: IPR002212; Fibril-assoc.
DR Pfam: PF00008; EGF; 45.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGF_BLOOD.
DR SMART: SMD0179; EGF_Ca; 43.
DR PROSITE: PS000010; ASX_HYDROXYL; 43.
DR PROSITE: PS000022; EGF-1; 2.
DR PROSITE: PS000187; EGF_Ca; 42.
DR PROSITE: PS000187; Calcium-binding; Glycoprotein; EGF-like domain;
KW Extracellular matrix; Calcium-binding; Disease mutation; Polymorphism.
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
FT SIGNAL 1 28
FT CHAN 29 2911
FT DOMAIN 11 142
FT DOMAIN 114 176
FT DOMAIN 145 176
FT DOMAIN 176 207
FT DOMAIN 275 316
FT DOMAIN 317 358
FT REPEAT 355 425
FT DOMAIN 493 533
FT DOMAIN 534 573
FT DOMAIN 615 615
FT DOMAIN 616 656
FT DOMAIN 657 697
FT REPEAT 698 766
FT DOMAIN 767 808
FT DOMAIN 809 850
FT DOMAIN 951 990
FT DOMAIN 954 995
FT REPEAT 996 1071
FT DOMAIN 1072 1113
FT DOMAIN 1114 1156
FT DOMAIN 1157 1198
FT DOMAIN 1199 1240
FT DOMAIN 1241 1281
FT DOMAIN 1282 1323
FT DOMAIN 1324 1365
FT DOMAIN 1366 1406
FT DOMAIN 1407 1447
FT DOMAIN 1448 1489
FT DOMAIN 1490 1530
FT DOMAIN 1531 1571
FT REPEAT 1572 1648

CC -----

FT DOMAIN 1649 1732
FT DOMAIN 1691 1732
FT REPEAT 1733 1806
FT DOMAIN 1807 1848
FT DOMAIN 1849 1890
FT DOMAIN 1891 1932
FT DOMAIN 1933 1974
FT DOMAIN 1972 2014
FT DOMAIN 2015 2054
FT DOMAIN 2055 2096
FT REPEAT 2097 2169
FT DCM1N 2170 2211
FT DCM1N 2212 2251
FT DCM1N 2252 2292
FT DOMAIN 2293 2336
FT DOMAIN 2337 2378
FT REPEAT 2379 2447
FT DOMAIN 2443 2489
FT DOMAIN 2490 2530
FT DOMAIN 2531 2569
FT DOMAIN 2570 2612
FT DOMAIN 2613 2652
FT DOMAIN 2653 2693
FT DISULFID 2694 2733
FT DISULFID 115 124
FT DISULFID 119 124
FT DISULFID 132 141
FT DISULFID 149 159
FT DISULFID 153 164
FT DISULFID 166 175
FT DISULFID 180 190
FT DISULFID 184 195
FT DISULFID 187 195
FT DISULFID 197 206
FT DISULFID 275 291
FT DISULFID 302 306
FT DISULFID 321 333
FT DISULFID 328 342
FT DISULFID 344 357
FT DISULFID 497 509
FT DISULFID 504 518
FT DISULFID 520 522
FT DISULFID 538 548
FT DISULFID 543 557
FT DISULFID 559 572
FT DISULFID 573 590
FT DISULFID 585 599
FT DISULFID 601 614
FT DISULFID 620 631
FT DISULFID 626 640
FT DISULFID 778 792
FT DISULFID 794 807
FT DISULFID 813 825
FT DISULFID 820 834
FT DISULFID 865 866
FT DISULFID 876 889
FT DISULFID 958 970
FT DISULFID 965 979
FT DISULFID 981 994
FT DISULFID 1076 1088
FT DISULFID 1083 1097
FT DISULFID 1109 1112
FT DISULFID 1118 1130
FT DISULFID 1125 1139
FT DISULFID 1141 1155
FT DISULFID 1161 1173

Query Match 13.0%; Score 727; DB 1; Length 2911;
 Best Local Similarity 26.5%; Pred. No. 240;
 Matches 255; Conservative 94; Mismatches 301; Indels 312; Gaps 51;

Qy 45 DVDECAGLDDPHADALCQNTPRSPYKSGKRYQG--EGRGCEDDECGNBLNGCVHDC 102
 Db 1199 DINECSLS-DNLCRNGKCNMIGTYQCSCNCNGYQATPDRQCTD-DEC-MINNGCDTQC 1256

Qy 103 LNPGNYRCTCFDGFMLAHDGHNCJLVDCELENN----GCCQHTCVNYNGSYECCCKEG 157
 Db 1257 INSEGSTECSSEGYALMPDERSCAIDEC-BNNPPDICDG---OCTNPGEXCLCYDS 1332

Qy 158 FFLSDNOHTCHRSEEGLSCNMKDHGSHICK---EAPRGSSVACECRPGFELAKNQRDC 213
 Db 1313 PMASMDMKTC-DVNEDLN-----SNICMFGESCENTRSFICCHGCGSYVKKGCTG 2364

Qy 214 -- ILTCNPHGNGC--OHSCDDTDATGPECSCHPCKYKHTBGSCLRCRDLTEVTEISNTS 269
 Db 1365 TDVDBEIGAHNCMDHASCLNIPGSFKCSREGW-IGNIKCIDLDEC----SNGT- 1415

Qy 270 VVDGDKERVKRRLIMETGAVNNGGCDRTCKDFTSTGVHCSCPYGETLQLDGKTKRDLDECQT 329
 Db 1416 -----HQCSIN----AOCVNPTRGSGTACSEGF--GJGTSIDVECAE 1454

Qy 330 RNGGCDH-FCKNIVGSPDGCKGKGFKLTDTEKSCDPOVDECSLDRTC-DHSCINHRGTFAC 397
 Db 1455 NINLCENQCLNPGAVRCECMGFTPASDERSQCIDECFQNCVSGTNUNLPGMFHC 1514

Qy 388 ACNRGFTL-YGTHCQDTNECS-IN-NGQQQCVQCVNTGVSYECOCHPGKWLHNWKDC 442
 Db 1515 ICDGVELDRTGGNCTDIDEADPINCVNG---LCVNTPSEYECNCPEPQLNPTGVG 1573

Qy 443 VEVK-GLLPTPSYSPRVLHCKSGGGDGCFLRCHSGIHLSSDVTIRTSTYFKNLNGKGS 501
 Db 1571 VDNVRGNCYLKGPR----GDGS-JLSCNTET---GVGVRSRSSCC-----CS 1608

Qy 502 LKNAELFPEGRPLPALPKHSSVYESFRYVNUUTCSSGKQVPAGPGRPSTPKEMF-TVEFEL 561
 Db 1609 LGRAWGNPCETCPV----NSTIYYTL-----CPGEGERPVPN----TTIJJED 1649

Qy 562 ETNQKEWTASCD-LSCD-VKRTEKRLKAIRTRKAYREQFHQLQESGMN:DVAKKPPRTS 620
 Db 1650 IDECOELPGLGCGNCI-----N 1667

Qy 621 EROAESCGVGQH-AE-----NOCVS---CRAGTYDGAERCLLQP----- 650
 Db 1668 TFGSFQECPQYIYLSDTRICEDECFAHPGVCGPCTCMTLGNYTCICPPEYMQVNG 1727

Qy 659 -----NG-TFONE-----EGONTCEPCEPGRPNSG----- 661
 Db 1728 GHNCMDMRKSFFYRSYNTGTCNCNELPNVTMRMCCTYAVKGAKNPCKPPTGTADEK 1787

Qy 682 -----ALKTPPEAMNNSEC----- 694
 Db 1788 TICGNIPGFTFDIHTGKAVDIDECKEKEPIGICANGVCINQGSFRCBPCTGFSYNLDEYC 1847

Qy 695 -----GGLCO-----PGEYSAJDGPAPCQLCALG-TFQPEAG--RTSCEFP--- 730
 Db 1848 EDIDECSNGNDLNCQRNADCINPSGRCE-----CAAGFKLNSNGACVDRNELEIF 1900

Qy 731 -CGGLATKXHQAATSFQ-DCETRVOQS PGHFYNTT-THR CIR--CPVGTQPEFGKRNCVS 726
 Db 1901 VCSHGLCVLDLG- SYQCICHHNGFKASQDQTMCMYDVECEHPCGNGTCRNTVGSYNCNL 1958

Qy 787 CPG-----NTTIDFDGSTNI--TOCKNRRGEGELGDF----TGYESPNYPGNYPA NT 833
 Db 1959 YPGFEL-THNNDCLDIDECSSFFQVCHNGRCFNEIGSFKCLCNEGYELTPDGKNCIDTN- 2017

Qy 834 EC 935
 Db 2018 EC 2019

RESULT 2
 FBX2_MOUSE STANDARD; PRT; 2907 AA.
 ID FBX2_MOUSE; STANDARD; PRT; 2907 AA.
 AC Q61555; Q63057;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE FBX2 OR FBX-2 precursor.
 GN FBX2 OR FBX-2.
 CS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TAXID=10090.
 PPI []
 SEQUENCE FROM N.A.
 MEDLINE=95203670; PubMed=7744363;
 RX RA "Developmental expression of fibrillin genes suggests heterogeneity of extracellular microfibris."
 RT J. Cell Biol. 129:1165-1176 (1995).
 RN 2.
 RP SEQUENCE CF 210-317 FROM N.A.
 RX MEDLINE=9410368; PubMed=8307578;
 RA Li X., Pereira L., Zhang H., Sangnireti C., Ramirez F., Bonadio J., Francke U.;
 RA "Fibrillin genes map to regions of conserved mouse/human synteny on mouse chromosomes 2 and 18." Genomics 18:667-672 (1993).
 CC FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS THAT BINDS COLLAGEN. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC SIMILARITY: Contains 47 EG-like domains.
 CC SIMILARITY: Contains 7 TGF beta binding protein (TGFBP) domains.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@iso-sib.ch.
 CC
 EMBL; I3979C; AAH74903; 1; -
 DR PIR; A57278; AAG6C695; 1; -
 DR HSSP; B35555; 1EMN.
 DR NGD; YGI:95349C; Pbr2.
 DR GO:GO-030326; P1imb morphogenesis; IMP.
 DR InterPro; IPR00152; Asx hydroxyL.
 DR InterPro; IPR01981; EGF-1.
 DR InterPro; IPR01338; EGF-1.
 DR InterPro; IPR005209; EGF-like.
 DR Pfam; PF00308; EGF; 45.
 DR PRINTS; PR00010; EGFBL00D
 SMART; SM00119; EGF CA; 43.
 DR PROSITE; PS00010; ASX HYDROXYL; 43.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 36.
 DR PROSITE; PS01187; EGF CA; 43.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Repeat; Signal; Multigene family.
 FT SIGNAL : 28 POTENTIAL.
 FT CHAIN : 29 FIBRILLIN.
 FT DOMAIN : 11: 4:2 EGF-LIKE 1.
 FT DOMAIN : 145: 176 EGF-LIKE 2.
 FT DOMAIN : 176: 208 EGF-LIKE 3.
 FT DOMAIN : 276: 317 EGF-LIKE 4, CALCIUM-BINDING.
 FT DOMAIN : 318: 359 EGF-LIKE 5, CALCIUM-BINDING.
 REPEAT : 360: 426 TGFBP 1.
 DOMAIN : 487: 527 EGF-LIKE 6.

EGF-LIKE 7, CALCIUM-BINDING.
EGF-LIKE 8, CALCIUM-BINDING.
EGF-LIKE 9, CALCIUM-BINDING.
EGF-LIKE 10, CALCIUM-BINDING.
TCFBP 2.
EGF-LIKE 11, CALCIUM-BINDING.
EGF-LIKE 12, CALCIUM-BINDING.
EGF-LIKE 13, CALCIUM-BINDING.
EGF-LIKE 14, CALCIUM-BINDING.
TCFBP 3.
EGF-LIKE 15, CALCIUM-BINDING.
EGF-LIKE 16, CALCIUM-BINDING.
EGF-LIKE 17, CALCIUM-BINDING.
EGF-LIKE 18, CALCIUM-BINDING.
EGF-LIKE 19, CALCIUM-BINDING.
EGF-LIKE 20, CALCIUM-BINDING.
EGF-LIKE 21, CALCIUM-BINDING.
EGF-LIKE 22, CALCIUM-BINDING.
EGF-LIKE 24, CALCIUM-BINDING.
EGF-LIKE 25, CALCIUM-BINDING.
EGF-LIKE 26, CALCIUM-BINDING.
TCFBP 4.
EGF-LIKE 27, CALCIUM-BINDING.
EGF-LIKE 28, CALCIUM-BINDING.
TCFBP 5.
EGF-LIKE 29, CALCIUM-BINDING.
EGF-LIKE 30, CALCIUM-BINDING.
EGF-LIKE 31, CALCIUM-BINDING.
EGF-LIKE 32, CALCIUM-BINDING.
EGF-LIKE 33, CALCIUM-BINDING.
EGF-LIKE 34, CALCIUM-BINDING.
EGF-LIKE 35, CALCIUM-BINDING.
TCFBP 6.
EGF-LIKE 36, CALCIUM-BINDING.
EGF-LIKE 41, CALCIUM-BINDING.
EGF-LIKE 42, CALCIUM-BINDING.
EGF-LIKE 43, CALCIUM-BINDING.
EGF-LIKE 48, CALCIUM-BINDING.
EGF-LIKE 49, CALCIUM-BINDING.
TCFBP 7.
EGF-LIKE 45, CALCIUM-BINDING.
EGF-LIKE 46, CALCIUM-BINDING.
EGF-LIKE 47, CALCIUM-BINDING.

DISULFID 528 636
DISULFID 569 655
DISULFID 650 675
DISULFID 691 690
DISULFID 692 777
DISULFID 760 786
DISULFID 761 788
DISULFID 802 801
DISULFID 803 801
DISULFID 844 819
DISULFID 845 839
DISULFID 893 854
DISULFID 948 870
DISULFID 989 870
DISULFID 990 870
DISULFID 1065 870
DISULFID 1066 870
DISULFID 1108 870
DISULFID 1150 870
DISULFID 1151 870
DISULFID 1192 870
DISULFID 1214 870
DISULFID 1235 870
DISULFID 1276 870
DISULFID 1317 870
DISULFID 1318 870
DISULFID 1359 870
DISULFID 1360 870
DISULFID 1400 870
DISULFID 1441 870
DISULFID 1442 870
DISULFID 1483 870
DISULFID 1484 870
DISULFID 1525 870
DISULFID 1565 870
DISULFID 1566 870
DISULFID 1643 870
DISULFID 1684 870
DISULFID 1727 870
DISULFID 1800 870
DISULFID 1842 870
DISULFID 1843 870
DISULFID 1885 870
DISULFID 1926 870
DISULFID 1927 870
DISULFID 1965 870
DISULFID 1966 870
DISULFID 2008 870
DISULFID 2048 870
DISULFID 2090 870
DISULFID 2091 870
DISULFID 2163 870
DISULFID 2205 870
DISULFID 2206 870
DISULFID 2245 870
DISULFID 2246 870
DISULFID 2286 870
DISULFID 2287 870
DISULFID 2330 870
DISULFID 2331 870
DISULFID 2332 870
DISULFID 2337 870
DISULFID 2441 870
DISULFID 2442 870
DISULFID 2483 870
DISULFID 2484 870
DISULFID 2524 870
DISULFID 2525 870
DISULFID 2563 870
DISULFID 2606 870
DISULFID 2607 870
DISULFID 2647 870
DISULFID 2688 870
DISULFID 2727 870
DISULFID 115 870
DISULFID 119 870
DISULFID 130 870
DISULFID 132 870
DISULFID 141 870
DISULFID 149 870
DISULFID 159 870
DISULFID 153 870
DISULFID 164 870
DISULFID 166 870
DISULFID 175 870
DISULFID 180 870
DISULFID 184 870
DISULFID 198 870
DISULFID 207 870
DISULFID 280 870
DISULFID 287 870
DISULFID 301 870
DISULFID 303 870
DISULFID 316 870
DISULFID 322 870
DISULFID 334 870
DISULFID 343 870
DISULFID 345 870
DISULFID 491 870
DISULFID 498 870
DISULFID 514 870
DISULFID 532 870
DISULFID 537 870
DISULFID 553 870
DISULFID 556 870
DISULFID 572 870
DISULFID 579 870
DISULFID 593 870
DISULFID 608 870
DISULFID 614 870
DISULFID 620 870

DISULFID 636 669
DISULFID 655 666
DISULFID 661 675
DISULFID 677 690
DISULFID 765 777
DISULFID 777 786
DISULFID 788 801
DISULFID 801 819
DISULFID 828 828
DISULFID 830 843
DISULFID 849 859
DISULFID 854 868
DISULFID 870 893
DISULFID 870 903
DISULFID 870 916
DISULFID 870 924
DISULFID 870 933
DISULFID 870 949
DISULFID 870 959
DISULFID 875 964
DISULFID 875 973
DISULFID 875 988
DISULFID 1070 1082
DISULFID 1077 1091
DISULFID 1106 1106
DISULFID 1112 1124
DISULFID 1119 1133
DISULFID 1135 1149
DISULFID 1155 1167
DISULFID 1162 1176
DISULFID 1178 1191
DISULFID 1197 1203
DISULFID 1204 1218
DISULFID 1220 1233
DISULFID 1239 1250
DISULFID 1246 1259
DISULFID 1261 1274
DISULFID 1280 1292
DISULFID 1287 1301
DISULFID 1303 1316
DISULFID 1322 1334
DISULFID 1329 1343
DISULFID 1345 1358
DISULFID 1364 1377
DISULFID 1371 1388
DISULFID 1388 1399
DISULFID 1405 1418
DISULFID 1412 1427
DISULFID 1429 1440
DISULFID 1446 1453
DISULFID 1463 1467
DISULFID 1463 1482
DISULFID 1468 1499
DISULFID 1494 1510
DISULFID 1510 1523
DISULFID 1529 1540
DISULFID 1535 1543
DISULFID 1551 1564
DISULFID 1647 1659
DISULFID 1654 1663
DISULFID 1670 1683
DISULFID 1689 1701
DISULFID 1696 1710

Query Match: Score 724; DB 1; Length: 2907;
Best Local Similarity: 26.1%; Pred. No. 3.e-40;
Matches 251; Conservative Matches 300; Indels 314; Gaps 51;

OY 45 DVDECAQDSDQHADLJCNQPTSKCSCKPGVQG--EGROCEDIDECENLNGCVCNHC-102
OY 1193 DINECSLS-CNLRNGKVNMGTYQCSNPGYQATPDRGQCTDIEC-MIMNGSCDCC 1250
OY :03 :NIPGNYRCTCQGFMALHDGHNCLDVEDECLENN---GCGQHCVNNGSYBCCKEG 157
OY 1251 TNSEGSTCS-SEGAMPDERSACDIEC-ENNDICG--QTNPGYROLCTG 1356
CY 158 FFLSDNHTC-HRSEEGLSYNKDHGSCHICK---EAPPGSVACECRGPELAKNORDC 21;
Do 13c7 ENASMDXHTC-DVNECDLN-----ENICMFGECENTGSFICHQCLQSYRKGTG 1358

214 -ILTCMHNGGC- QHSQDDTADPECSCHPQYKMHHTDGRSCLEREDTVJLETESNTTS 269
 QY :|:||| :|:||| :|:||| :|:||| :|:
 Db 1359 TDVDECEIGAPNCMDHASCUNPGFCKSCREGVW-GNSIKCIDLDEC----ANGI- 1409
 Qy :|:||| :|:||| :|:||| :|:
 Db 270 VVDGDKVRKRLLMETCAVNNGGCDRTCKD-TSTGYHCSEGVFTQLDQTKRDKDIDECOT 329
 Qy :|:||| :|:||| :|:
 Db 1410 ------HOCSIN----AQCVNTPGSYRCACSEGFT-GDGFTCSVDCEAE 1448
 Qy :|:||| :|:||| :|:
 Db 333 RONGCDH-FCHNIVGSFDGCCKGKPKLTLDEKSQDVDECSLORTCDH-SCINHPGTFAC 387
 Qy :|:||| :|:||| :|:
 Db 1449 MTNLCENGQCJNVPGAVRCBEMGFTPASDRSQRSCDIDEFSQNITCVEGTNN-EGMPHIC 1508
 Qy :|:||| :|:
 Db 388 ACNRGVTI-YGFTHCGDTNECS-IN-NSGQQCQVNYNTGSYEQCHRGYKLHWNKOC 442
 Qy :|:||| :|:
 Db 1509 ICDDGFELDRGKNTDIEADPINCNGV---LCVNTRGTRYEBNCPPQUNATGVC 1564
 Qy :|:||| :|:
 Db 443 VEVK-GLLPTVSPrVSLHCGSGGGDFCFRCHSGCZHLSDDVTIRTSVTFKLNEGKGS 501
 Qy :|:||| :|:
 Db 1565 UDNRVGNCLKFGRPR-----GDGS-LSCNT---EAGVGYSRSSCC----CS 1632
 Qy :|:||| :|:
 Db 502 IKNAEFLP-BGLRPALPEKHSSVKESFRYNLTCSSGKQVPGAPGRSPTPKEMFITVERE 560
 Qy :|:||| :|:
 Db 1603 IASKAWNPCTCOPVNSTEHTYL-----CPGGEFRPNP---ITIE 1642
 Qy :|:||| :|:
 Db 561 LETNQEVTASCD-LSCIVKTEKRURKA.FT.LRKAVHRECFLQLSGMNDVAKKPPR. 6.9
 Qy :|:||| :|:
 Db 1643 DIDEQCLPLPGLGQGGNNI-----
 Qy :|:||| :|:
 Db 620 SERQAESCGVGCGH-AE-----NQCVS----CRAGTYKDGAERCILCP---- 1660
 Qy :|:||| :|:
 Db 1661 NTFGSFQCECPGQYYLSEETRICEDECFAHPGVGPCTGNTLGNYTCICPEYQVN 1720
 Qy :|:||| :|:
 Db 659 -----NG-TFQNE-----EGQMTCEPCPRPGNSG- 681
 Qy :|:||| :|:
 Db 1721 GHHCNCDMRKSFCCYRSYNGTICENELPFNVTKRMCCCTVWKGAGNKPCBPPGTADF 1783
 Qy :|:||| :|:
 Db 682 -----ALKPEAMNMSEC----- 694
 Qy :|:||| :|:
 Db 1781 KTGIGNIPGFFEDIHGRAVIDIDEKEIPGICANGVYCINQISFRCECPTGFSYNDLLY 184C
 Qy :|:||| :|:
 Db 695 -----GGLCQ-----PGEYSADGFAPCOLCAG-TPEOAG--RTSCFP-- 730
 Qy :|:||| :|:
 Db 1841 CEDIDECSNGDNLCORNADCINSPGSYRCE-----CAAGFKLSPNGACVDRNECLE-P 1893
 Qy :|:||| :|:
 Db 731 -CGGLATRKGATGATSQ-DETRVQSPGFIYNTTHRCR-R-CPGTYQPEFSRNNCV 785
 Qy :|:||| :|:
 Db 1894 NYCSHGLCVDLQG--SYQCIQNGNFRASQDTMCDVDECRHPHCPANGTKNTGSYNC. 1951
 Qy :|:||| :|:
 Db 786 SCPG-----NTTDDFGSTNI---TQCKNRGGELGDF----TGYIESPNYPGNYFAN 832
 Qy :|:||| :|:
 Db 1952 CYPGFELTHNNDCLDIDECSSFFGQVCRNGRCFNEIGSFKLCNBEGYELTPDGKNCIDTN 2011
 Qy :|:||| :|:
 Db 833 TEC 835
 Qy :|:||| :|:
 Db 2012 -EC 2013

STRAIN=BALB/c; TISSUE=Limb;
 MEDLINE=97216965; PubMed=3083061;
 RA "primary structure and expression of matrixin-2, the closest relative
 of cartilage matrix protein within the von Willebrand factor type A-
 like module superfamily";
 RT J. Biol. Chem. 272:9268-9274 (1997);
 CC :- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
 CC :- SUBCELLULAR LOCATION: Secreted.
 CC :- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
 CC CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
 CC OSTEOBLAST CELL LINES.
 CC :- SIMILARITY: Contains 10 EGF-like domains.
 CC :- SIMILARITY: Contains 2 VWF-A domains.
 CC :-
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC :-
 ENBL: U69262; AAC53163.;
 DR M61:103613; Matr2
 DR HSSP: B05039; IAQ5.
 DR GO:0005578; C:extracellular matrix; IDA.
 DR InterPro: IPR000152; Axin hydroxyl.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR006239; EGF-like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR003035; VWF-A.
 DR PFM: PR03008; EGF; 10.
 DR PFM: PR03008; EGF; 2.
 DR PRINTS: PR00453; VWF DOMAIN.
 DR SMART: SMC0_81; EGF; 10.
 DR SMART: SM0327; VWA; 2.
 DR PROSITE: PS00010; AXIN HYDROXYL; 7.
 DR PROSITE: PS0234; VWF; 2.
 DR PROSITE: PS0234; VWF; 9.
 KW EGFL-like domain; Sigma; Glycocalyx; Repeat; Coiled coil.
 FT SIGNAL 1 23
 FT CHAIN 24 956
 FT DOMAIN 57 212
 FT DOMAIN 238 219
 FT DOMAIN 279 319
 FT DOMAIN 320 361
 FT DOMAIN 361 402
 FT DOMAIN 443 483
 FT DOMAIN 484 524
 FT DOMAIN 525 565
 FT DOMAIN 566 605
 FT DOMAIN 607 647
 FT DOMAIN 655 830
 FT DOMAIN 917 955
 FT DISULFID 242 253
 FT DISULFID 249 262
 FT DISULFID 264 277
 FT DISULFID 263 294
 FT DISULFID 290 303
 FT DISULFID 324 318
 FT DISULFID 331 344
 FT DISULFID 346 359
 FT DISULFID 365 376
 FT DISULFID 372 385
 FT DISULFID 387 400
 FT DISULFID 406 417
 FT DISULFID 426 426
 FT DISULFID 428 441
 FT DISULFID 447 458
 FT DISULFID 454 467
 RESULT 3
 MTN2_MOUSE STANDARD PRT: 956 AA.
 AC 008746;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Matrixin-2 (precursor)
 GN MTN2.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mls.
 OX NCBI_TaxID=10393;
 RN [1] _SEQUENCE FROM N.A.
 RP

FT	DISULFID	469	482	BY SIMILARITY.
FT	DISULFID	498	499	BY SIMILARITY.
FT	DISULFID	495	508	BY SIMILARITY.
FT	DISULFID	510	523	BY SIMILARITY.
FT	DISULFID	529	540	BY SIMILARITY.
FT	DISULFID	536	549	BY SIMILARITY.
FT	DISULFID	551	564	BY SIMILARITY.
FT	DISULFID	570	581	BY SIMILARITY.
FT	DISULFID	577	590	BY SIMILARITY.
FT	DISULFID	592	605	BY SIMILARITY.
FT	DISULFID	611	622	BY SIMILARITY.
FT	DISULFID	618	631	BY SIMILARITY.
FT	DISULFID	633	646	BY SIMILARITY.
FT	CARBONYD	221	221	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	890	890	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	956 AA;	106779 MW;	3BAC22770B66EEE CRC64;
Qy	Query Match	12.3%	Score 6.92;	DB 1; Length 956;
Matches 159; Conservative 29.8%; Fred. No. 1.3e-18;	Indels 78; Gaps 14;			
Db	62 CNTPTSYKCSCKPGY--QEGRQEDIDEQGNEELNGCVHDNUNTPGNYRCTFDGFNL	119		
Db	253 CLNTPTSYIICKCKQGYLSTDKTCKTQIQLCATE DHGECQLCVNMLGSFVQQCGSYTL	311		
Qy	120 AHGDHGNLDVDECLENNGGCQHTCYNVMGSYECCKEGFFLSDNOHTCIRHSEEGLSCHNN	179		
Db	312 AEDGKRATMDCASENHNHCHECYNAESSYLRCHEGFAALNSDKXTCISKID---YCAS	367		
Qy	180 KDHGCSHICKRAPIRGSVACEPRGPFLAKNRDC--ILTCHNGNGCCSHCDTADGPFC	237		
Db	368 SNHGCGOHECYNA-QTALCRCLKGPMLNPDRTKCRINRYCALNKNGPGEIEBCVNTEEGHYC	425		
Qy	238 SCHPQYKMHHTGRSCLE-----REDTVLETEVSNTTSVNG---DKRVVKRLLME	284		
Db	427 SCKROGNLDPLRGKTCRSVHDCAQODDHGCQEQLNTEESFYQCSEGFPLINDLKTCSRAD	465		
Qy	285 ?CAVNNGGCDTCKDSTGVKSCPVGFTLQJDGTICRKD?DECQTRNGGDDHFKRNIVS	344		
Db	487 YCLLSNHGCESYCVNTDKSFPCQQCPEGHVLRSDGKTCALKDSCALGDHSCHEHCVSSDDS	546		
Qy	345 FDGCKRKKGPKLTDKEKSQDVEDE-S-DRTQDHSCINHPGT-FACACNRGTYL-YGFTHCG	402		
Db	547 FVCQCFEGYIIRDGKTCRNDVQDQVNHGCCQDGLVNSGSESYVERCL2GFRLAEDGKRR	506		
Qy	403 D?NECSINNGGCOQCVNTVSYE?COPHCGYKLHKNKKOCVE-----	444		
Db	607 RKNVCKSTQHSE-HMCVNNGNSYLCRCSSEGVLIAEGK;CRCTEGSPID;VFVIDGSKS-	666		
Qy	445 -----VKGLLP?-SWSPRVSLHCKSGGGDGCFLRCHSGHILSSDVTTIRTST?	492		
Db	667 GREENFETVKHVTGILDS-LAVSKA-----ARVGLIQLYSTQVRTERT	703		
Qy	493 FK---INEGRCSLKAELFPLGLRPALEKSYKESFRYVNLTCSSGKVP 54;			
Db	709 LRGFSSAKEMVKAVTHMKYMGKSMGTMGLALK-MFERSFTQVEGARPPSTQF 760			
Qy	445 -----VKGLLP?-SWSPRVSLHCKSGGGDGCFLRCHSGHILSSDVTTIRTST?	492		
Db	667 GREENFETVKHVTGILDS-LAVSKA-----ARVGLIQLYSTQVRTERT	703		
Qy	493 FK---INEGRCSLKAELFPLGLRPALEKSYKESFRYVNLTCSSGKVP 54;			
Db	709 LRGFSSAKEMVKAVTHMKYMGKSMGTMGLALK-MFERSFTQVEGARPPSTQF 760			
AC	P98133;			
AC	P98133; BOVIN	STANDARD;	PRT:	2871 AA.
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DE	Fibrillin 1 precursor (MP340)			
GN	FBNI.			
OS	Bos taurinus (Bovine)			
OC	Eukaryota; Metacoda; Chordata; Craniata; Vertebrata; Eutelostomii; Mammalia; Buterida; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;			
OC	Bovidae; Bovinae; Bos.			
OX	[1]			
RN				

RESULT 4
 FBNI_BOVIN
 ID FBNI_BOVIN
 STANDARD:
 PRT: 2871 AA.
 AC P98133;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE Fibrillin 1 precursor (MP340)
 GN FBNI.
 OS Bos taurinus (Bovine)
 OC Eukaryota; Metacoda; Chordata; Craniata; Vertebrata; Eutelostomii; Mammalia; Buterida; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX [1]
 RN 1.97 1237

RP	SEQUENCE FROM N.A.			
RC	TISSUE-SKIN;			
RX	MEDLINE=95117592; PubMed=7835900;			
RA	Tilstra D.J., Potter K.A., Byrs P.H.,			
RT	"Sequence of the coding region of the bovine fibrillin cDNA and localization to bovine chromosome 1C."			
RR	Genomics 23:480-485(1994);			
RL	12;			
RN				
RP	PARTIAL SEQUENCE.			
RX	MEDLINE=96132851; PubMed=8557636;			
RA	Gibson M.A., Hatzinikolas G., Kumariatiakae J.S., Sandberg L.B., Nicholl J.K., Sutherland G.R., Clary E.G.,			
RC	"Further characterization of proteins associated with elastic fiber microfibrils including the molecular cloning of XAGP-2 (M25);"			
RT	J. Biol. Chem. 271:10956-10961 (1996).			
CC	- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS			
CC	THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE LONG-TERM FORCE BEARING STRUCTURAL SUPPORT			
CC	- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE MICROFIBRILS			
CC	- SIMILARITY: Contains 7 TGF-beta banding protein (TGFBP) domains.			
CC	- SIMILARITY: Contains 7 TGF-beta banding protein (TGFBP) domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb-sib.ch/announce/ or send an email to licensee@isdb-sib.ch).			
CC	MICROFIBRILS			
CC	- SIMILARITY: Contains 7 TGF-beta banding protein (TGFBP) domains.			
CC	DR: I28748; AAA74-22.1; -			
CC	DR: P15567; A55567.			
CC	DR: P35557; 1APJ.			
CC	DR: InterPro: IPR001512; Asx hydroxyl.			
CC	DR: InterPro: IPR001512; Asx hydroxyl.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: InterPro: IPR0529; EGF_Ca.			
CC	DR: PROSITE: PS000012; ASX_HYDROXYL; 43.			
CC	DR: PROSITE: PS000012; ASX_HYDROXYL; 43.			
CC	DR: PROSITE: PS01186; EGF_2; 38.			
CC	DR: PROSITE: PS01187; EGF_Ca; 43.			
CC	KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Signal; Multigene family.			
FT	SIGNAL 1 27 POTENTIAL.			
FT	CHAIN 28 2871 FIBRILLIN 1.			
FT	DOMAIN 81 112 EGF-LIKE 1.			
FT	DOMAIN 115 146 EGF-LIKE 2.			
FT	DOMAIN 147 178 EGF-LIKE 3.			
FT	DOMAIN 246 287 EGF-LIKE 4.			
FT	DOMAIN 288 329 EGF-LIKE 5.			
FT	REPEAT 330 390 CALCIUM-BINDING.			
FT	REPEAT 392 446 PRO-RICH.			
FT	DOMAIN 449 489 EGF-LIKE 6.			
FT	DOMAIN 530 571 EGF-LIKE 7.			
FT	DOMAIN 572 612 EGF-LIKE 8.			
FT	DOMAIN 613 653 EGF-LIKE 9.			
FT	REPEAT 654 722 CALCIUM-BINDING.			
FT	DOMAIN 723 764 EGF-LIKE 11.			
FT	DOMAIN 765 806 EGF-LIKE 12.			
FT	DOMAIN 807 846 EGF-LIKE 13.			
FT	DOMAIN 910 952 EGF-LIKE 14.			
FT	REPEAT 952 1227 TGFBP 3.			
FT	DOMAIN 1028 1369 EGF-LIKE 15.			
FT	DOMAIN 1070 1112 EGF-LIKE 16.			
FT	DOMAIN 1113 1154 EGF-LIKE 17.			
FT	DOMAIN 1155 1196 EGF-LIKE 18.			
FT	DOMAIN 1197 1237 EGF-LIKE 19.			

- RP REVIEW ON VARIANTS.
RX MEDLINE=98062175; PubMed=9401003;
RA "Fibrillin-1 mutations in Marfan syndrome and other type-1 fibrillinopathies." ;
RT "Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene." ;
RL Hum. Mutat. 10:415-423(1997).
RN [12]
RP VARIANT MFS PRO-1137.
RX MEDLINE=94130456; PubMed=1852208;
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y., Corson G.M., Puffenberger E.G., Hanosh A., Nanthakumar E.J., Cristianini S.M., Steet G., Meyers D.A., Francomano C.A.;
RT "Clustering of fibrillin (FBNI) missense mutations in Marfan patients at cysteine residues in EGF-like domains." ;
RL Hum. Mutat. 1:366-374(1992).
RN [13]
RP VARIANT MPS SER-1249; ARG-1663; SER-2221 AND SER-2307.
RX MEDLINE=94125844; PubMed=1301446;
RA Dietz H.C., Saraiya J.M., Pyeritz R.E., Cutting G.R., Francomano C.A., Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;
RT "Marfan phenotype variability in a family segregating a missense mutation in the epidermal growth factor-like motif of the fibrillin gene." ;
RL J. Clin. Invest. 89:1674-1680(1992).
RN [14]
RP VARIANT MFS ILI-548 AND ALA-723.
RX MEDLINE=98010946; PubMed=8406197;
RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C., Pyeritz R.E., Francomano C.A.;
RT "Four novel FBNI mutations: significance for mutant transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome." ;
RL Genomics 17:468-475(1993).
RN [15]
RP VARIANT MFS SER-2114.
RX MEDLINE=94278402; PubMed=8504310;
RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;
RT "A novel fibrillin mutation in the Marfan syndrome which could disrupt calcium binding of the epidermal growth factor-like module." ;
RL Hum. Mol. Genet. 2:475-477(1993).
RP VARIANT MFS ARG-862; TYR-1117; PRO-1137 AND PHE-1589, AND VARIANT ALA-1148.
RX MEDLINE=94108431; PubMed=8281141;
RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C., Berg M.A., Miller D.C., Francke U.;
RT "Mutation screening of complete fibrillin-1 coding sequence: report of five new mutations, including two in 8-cysteine domains." ;
RL Hum. Mol. Genet. 2:1813-1821(1993).
RN [16]
RP VARIANT MFS GLY-217 AND ARG-2627.
RX MEDLINE=95067970; PubMed=7977166;
RA Karttunen L., Raghunath M., Loennqvist L., Peltonen L.;
RT "A compound-heterozygous Marfan patient: two defective fibrillin alleles result in a lethal phenotype." ;
RL Am. J. Hum. Genet. 55:1083-1091(1994).
RN [17]
RP VARIANT EL YLS-2447.
RX MEDLINE=94245249; PubMed=8188302;
RA Loennqvist L., Child A., Kainulainen K., Davidsson S., Puhakka L., Peltonen L.;
RT "A novel mutation of the fibrillin gene causing ectopia lentis." ;
RL Genomics 19:573-576(1994).
RN [20]
RP VARIANT MFS CYS-627.
RX MEDLINE=94272497; PubMed=800412;
- RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;
RT "Two novel mutations and a neutral polymorphism in EGF-like domains of the fibrillin gene (FBNI): SSCP screening of exons 15-21 in Marfan syndrome patients." ;
RL Hum. Mol. Genet. 3:373-375(1994).
RN [21]
RP VARIANT MFS CYS-122.
RX MEDLINE=94314977; PubMed=8040326;
RA Stahl-Bärlund C., Ukkonen T., Kainulainen K., Kristoffersson U., Saxe T., Torngqvist K., Peltonen L.;
RT "An extra cysteine in one of the non-calcium-binding epidermal growth factor-like motifs of the FBNI polypeptide is connected to a novel variant of Marfan syndrome." ;
RL J. Clin. Invest. 94:709-713(1994).
RN [22]
RP VARIANT MFS TYR-1223.
RX MEDLINE=94351682; PubMed=8071963;
RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;
RT "A new missense mutation of fibrillin in a patient with Marfan syndrome." ;
RL J. Med. Genet. 31:338-339(1994).
RN [23]
RP VARIANT XES HIS-117C.
RX MEDLINE=95174777; PubMed=7870075;
RA Hayward C., Porteous M.E.M., Brock L.J.;
RT "A novel mutation in the fibrillin gene (FBNI) in familial atrachnodyctyly." ;
RL Mol. Cell. Probes 8:325-327(1994).
RN [24]
RP VARIANT MFS GLY-217; ASN-1223; ARG-1074; TYR-1242; ARG-1533;
GLY-2117; TRP-2251; LYS-2447 AND ARG-2511.
RX MEDLINE=94184368; PubMed=8136837;
RA Kainulainen K., Puhakka L., Sakai L., Peltonen L.;
RT "Mutations in the fibrillin gene responsible for dominant ectopia lentis and neonatal Marfan syndrome." ;
RL Nat. Genet. 6:64-69 (1994).
RN [25]
- Query Match Score 691; DB 1; Length: 2871;
Best Local Similarity 22.5%; Pred. No. 5e-38; Indels 442; Gaps 45;
Matches 269; Conservative 105; Mistratches 378; Index 378;
- QY 45 DVDECAQCLDCHADALCQNTPTSYKCSCKPASYQEGRCQEDIDEBCNELNGGCYEDCLN 104
DB 2363 DLDECNSHTKQSQHADDNTGSIYRLLCKESEYGGFTCDLDESERNLCGNGSQLN 1422
QY 105 *PONYRCFCDFGMLAHGHNCLVDCELENNGGCCTVNYMGSYECCKEGFF-SDNC 164
DB 1423 AFGYRCEDWGMGPSPAGKRAECDISLPRICVFCHNPGLRCECEIGYELDJSQ 1492
QY 165 HCIIHRSEBGSOMXKQHGCSHICKEARPGSVAECPGEFLAKNQEDCIL- 216
DB 1483 GNCTDNNE---CLDPPTCISRCNTP-GSYICDCPPDFELNP?PRVGCVDTSGNCYLD 1537
QY 217 -----
DB 1538 IRRPGDNQDTASNEIGYGUSKACCCSICKANGTCEMCRAVNTSEYKILCPGGEGFRP 1597
QY 217 -----CNHGNNGCC-----HSCTDTAD-----HSCTDTAD-----233
DB 1598 NP-TVILEDIDECQELGQCGKCNTFGSPQCRCPTGYJNEDTRVCDVNNECTP3G- 1657
QY 234 -GP-----ECSCHEQKHTDGRSCLEREDYVELETSNTSVWDGDKRYKRLL 292
Do 1658 CGGSTCYNTVGNYCICPPDQMNGNQCYMRSCLCYRNYYADNQTCDCGELENMTKK 1717
QY 263 METGAVNS-GCDRICK-----DTSTGV-----304
DB 1718 MCCCSTYNGKGRANMKPQQCIPFESTDFATLGSSQRDGFVITYG-PFDIDSCRE_PGVVC 1777
QY 305 -----HCSCPVGFTIQLDGKTKOJIDECC-----328
DB 2773 ENGVCIINNGSFRCECPYGFYNQKLUVEDDECGPVCQNAECINTAGSYCNDCKP 1637

RC STRAIN=CD-1; TISSUE=Kidney;
 RA Cta K.; Kumar A.; Wada J.; Liu Z.; Karwar Y.S.;
 RL Submitted [APR-1995] to the EMBL/GenBank/DDBJ databases.
 CC --!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONNECTING MICROFIBRILS PROVIDE
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC --!- PREDICTED INTERACTIONS: FIBRILLIN-1 DISULFIDE BONDS EITHER WITH OTHER
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
 CC MICROFIBRILS (BY SIMILARITY).
 CC --!- SIMILARITY: Contains 7 TGF-beta binding protein (TGF β) domains.
 CC --!- SIMILARITY: Contains 7 EGF-like domains.
 CC --!- SIMILARITY: Contains 7 EGF-like domains.

This SWISS-PRC entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation: the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL: L29454; AAAS640_1; --.
 CC DR EMBL: AAS642_7_1; --.
 CC DR PIR: AAS649_1; AAAS640_1; --.
 CC DR HSSP: P35555; 1AFJ.
 CC DR InterPro: IPR001881; EGF_Ca.
 CC DR InterPro: IPR005209; EGF_Like.
 CC DR InterPro: IPR005212; Fibrillin_assoc.
 CC DR Pfam: PF00039; EGF; 46.
 CC DR Pfam: PF00683; FB; 9.
 CC DR SMART: SM0179; EGF_Ca; 42.
 CC DR PROSITE: PS00010; ASY_HYPERXY; 43.
 CC DR PROSITE: PS00322; EGF_1; 2.
 CC DR PROSITE: PS001186; EGF_2; 38.
 CC DR PROSITE: PS001187; EGF_Ca; 43.
 CC KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Repeat; Signal; Multigene family.
 CC FT SIGNAL 1 27
 CC FT CHAIN 28 2871
 CC FT DOMAIN 8 112
 CC FT DOMAIN 1:5 146
 CC FT DOMAIN 14:7 178
 CC FT DOMAIN 24:6 287
 CC FT DOMAIN 288 329
 CC FT DOMAIN 330 401
 CC FT DOMAIN 4:2 446
 CC FT DOMAIN 44:9 489
 CC FT DOMAIN 49:3 529
 CC FT DOMAIN 53:0 571
 CC FT DOMAIN 57:2 612
 CC FT DOMAIN 61:3 653
 CC FT DOMAIN 65:6 721
 CC FT DOMAIN 72:3 764
 CC FT DOMAIN 76:5 826
 CC FT DOMAIN 80:7 846
 CC FT DOMAIN 91:0 951
 CC FT DOMAIN 95:2 1018
 CC FT DOMAIN 102:8 1069
 CC FT DOMAIN 107:0 1112
 CC FT DOMAIN 111:3 1154
 CC FT DOMAIN 115:5 1196
 CC FT DOMAIN 123:8 1279
 CC FT DOMAIN 128:0 1321
 CC FT DOMAIN 132:2 1362
 CC FT DOMAIN 136:3 1403
 CC FT DOMAIN 140:4 1445
 CC FT DOMAIN 144:6 1486
 CC FT DOMAIN 148:7 1527
 CC FT DOMAIN 152:8 1599
 CC FT DOMAIN 160:6 1647
 CC FT DOMAIN 164:8 1688

RESULT 6

FBNI_MOUSE_ID_FBN1_MOUSE STANDARD; PRT; 2871_AA.
 AC Q61554; Q60826.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Fibrillin 1 precursor.
 GN FBN1 OR FBN-1.
 OS Mus musculus (Mouse); Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Sciurognathi; Muridae; Murine; Mus.
 DC Mammalia; Eutheria; Rodentia;
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9510561; PubMed=7829516;
 RA Yin W.; Germiller J.J.; Sanguineti C.; Smiley S.; Pangilinan T.,
 OC Pereira L.; Ramirez F.; Bonadio J.;
 RA "Primary structure and developmental expression of Fbn-1, the mouse
 OC fibrillin gene";
 RL J. Biol. Chem. 270:1798-1806(1995).
 RN SEQUENCE FROM N.A.

FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1221	1221	BY SIMILARITY.
FT	DISULFID	1223	1223	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DISULFID	1415	1423	BY SIMILARITY.
FT	DISULFID	1432	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	1658	1672	BY SIMILARITY.
FT	DISULFID	1674	1687	BY SIMILARITY.
Query Match		12.3%	Score 690; DB 1;	Length: 2871;
Best Local Similarity		23.4%	Pred. No. 5..8e-19;	
Matches	276;	Conservative	Indel's 418;	Gaps 51;
Qy	45 DVDECAQQLDCHADALCINTPTSYKCKPSYQGBRQCEDDECONELNGSCVHCCLN 164			
Db	:1363 DLDECNSGTHMCQSQADQNTNGSYRCLKDQSYTGDDFTCDLDECSENLMUGNGQCLN 1422			
Qy	105 IPGNYRCYTFDGFMLAHHGHNCDLVDGS-----:-----:-----:LENG 137			
Db	1423 APGYKRCZEDGMYPSACKRAEDEDLSLNPICUVECTOCHNLPGLEPCECEIGYELDRS 1482			
Qy	138 G-----:-----:CGETCIVANGMSYECCKKEGFF-SENGHTC:-----:RS 171			
Db	14B3 GNCTDVBNCJDPITC:SGNCVNTPGSY:TDCSPDFENLPTRVGCVDTSGNCYLNRPG 1542			
Qy	172 EEGLSCKMKQHG-----:-----:CSHL-----:-----:-----:-----:-----:187			
Db	1543 DNGTACNEIGGVSKASCCCSLGRKAWGTCBLCPSPVNTSEYKILCPGGEGFRENPTV :EC2			
Qy	188 -----:CKEAPR-----:-----:GSVACECPGPGLAKQNQDC--:ILTCHN-----:GNGG 223			
Db	1603 IIEDIDECGELPQJCGGKCIINTFGSSEQCRCTGYY:NETPRAVCDDBTLEVTENTTSVYGDGRKSRLLM 283			
Qy	224 CQHSCDDTADGPESCSHQYKQHHTDGRSLCRLBDTVELENTTSVYGDGRKSRLLM 283			
Db	1662 ---TCYNTVGNVYC:CPDYMQUNGNNCMQRSSIVYRNYYADNCTCDGELJFNMTKC: 1718			
Qy	284 ETCAVN-----:NGGCDRCKDST-----:-----:-----:-----:3G2			
Db	1719 CCCSYNIRRAKPKCQ.CPIB:TCFATLGCSQRPGFVID:YGLPVIDECREIPGVC 1777			
Qy	3C3 --GV-----:-----:HSCSPYGETLQJDGKTCRK:DECO-----:-----:328			
Db	1778 ENGCIAWNGSFCCECPGFYFNKLVCE:-----:DECONSPVCLRKAFCINTASYRDKP 1637			
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1043	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.

Qy	371	ARTCDHSCINPFGTACACNRYTYLGFTHGCGTNECSINRG-COO	WCVNTVGSYECQ	438	like module superfamily;
RN	13	J. Biol. Chem.	272:2268-9274 (1997).		
Db	1898	DAGNGTCRNTIGSFCNRNGFILSHNNODIVDECATNGNLCRNGCVENTGSFCQR	1957	RP	SEQUENCE 244-956 FRCM N.A.
Qy	429	CHPGYKLHNNRKDCEVKGK--LLPTPSVSPRLHCKSGGGDFLRCNSGHISLSDVTT	486	RC	TISSUE_Tessis;
Db	1958	CNEGYVAPDRTCDINECLDGRCAVT--CQNDGSYRCI--CPGYSIQNDICE	2012	RA	Submitted [JAN-2001] to the EMBL/GenBank/DBJ databases.
Qy	487	IRTSVTEFKLNK-----GKCSL-KNAE-----LPFEGLR-PALPERHSYKESFRYY	530	CC	- FUNCTION: "INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
Db	2013	-----DIDCVCVEEPICAGLCTS--NTBOSFKLCLPCEGFSWSSGRDQJRMXYCA	2063	CC	- SUBCELLULAR LOCATION: Secreted.
Qy	531	NLTCSGGKQVQVGPARGPRSTPKEMITFEELETFETQEKEVASCDS-----CIVKTRTEK	583	CC	- ALTERNATIVE PRODUCTS:
Db	2064	KF-----EGGICSSPK-----SRNSHKOECALLGEGMGDPCLCIPTEP	2103	CC	- Event: Alternative splicing; Named isoforms=2;
Qy	584	RFLRKAITRLRAVHRBFHQI-----SGMNLDAVKPP-----RTSER	622	CC	- Name=Long;
Db	2104	-----DEAFATCICPPGSGIYWPDOSAVDKECDECEPDPVDRHQGINTDGS	2148	CC	- isoID=000339-1; Sequence=VPS_0C1399;
Qy	623	QAESCGYQGHABNOVY---SRRAGTYDGARERCTCILCPNGTFCOPPRN	679	CC	- !- SIMILARITY: Contains 10 EGF-like domains.
Db	2149	YRCECPGYYILEGENCUTDCSVGN-----PCGNGTCRNVIGGFC-TCEEGEE	2197	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Qy	680	SGALKTPEAMMSECGG-----LCQGPYESADGFAPCOLCALGTFPEAER-SC----FPCGG	733	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Db	2198	#PMMPMC-----DNECAQNPLICAFCRVNTGTYSYECK-CPVSYVURE-DRMKCKDECAE	2252	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Qy	734	GATKHCQGATSFDQDCETRQ-----CSPGH-YNTITHRCIRCPCVSTYQCEFGKNCVSCP	768	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Db	2254	G----KHDCTEKOMECKMLNYMCICGPYORRDPGEGL-----DEBQCTkP	2233	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Qy	789	GNTTTDEGTTINITQCKNRRGCGELGDFP-----GYIESNYPGNYPANTECT-WTINAPP	843	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Db	2300	G-----ICENGRCNLNTGSYTCBENDGTAS-----PTQECI-----	2312	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Qy	844	KRRILIVVPEIFLPIEDDGDYLVMRKTTSSNSVTVETCTYER-----PIAFTS	934	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Db	2333	DNREGYCSEVF--ENMC-----QIGSSNNNPVTKSECCVGGKWLHLCEICPFEGTV	2384	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Qy	895	RSKK-WIQLKSNEGNSNARGFPQPVYTYDEDQELIEDVRS	936	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
Db	2365	AYRKLC-----PHGRGMNTNGADYDE--CKVHIDVCRNS	2416	CC	- !- SIMILARITY: Contains 2 VWF-A domains.
				XW	EGF-like domain: Signal; Glycoprotein; Repeat; Coiled coil;
				POTENTIAL.	
				SIGNAL	
			1	2	
			FT	CHAIN	
			24	24	MATRILIN-2.
			FT	DOMAIN	
			57	57	VWFA 1.
			FT	DOMAIN	
			238	278	EGF-LIKE 1.
			FT	DOMAIN	
			279	319	EGF-LIKE 2.
			FT	DOMAIN	
			320	360	EGF-LIKE 3.
			FT	DOMAIN	
			361	401	EGF-LIKE 4.
			FT	DOMAIN	
			402	442	EGF-LIKE 5.
			FT	DOMAIN	
			443	483	EGF-LIKE 6.
			FT	DOMAIN	
			484	524	EGF-LIKE 7.
			FT	DOMAIN	
			525	563	EGF-LIKE 8.
			FT	DOMAIN	
			566	606	EGF-LIKE 9.
			FT	DISULFID	
			276	276	BY SIMILARITY.
			FT	DISULFID	
			264	277	BY SIMILARITY.
			FT	DISULFID	
			283	294	BY SIMILARITY.
			FT	DISULFID	
			290	303	BY SIMILARITY.
			FT	DISULFID	
			305	318	BY SIMILARITY.
			FT	DISULFID	
			324	335	BY SIMILARITY.
			FT	DISULFID	
			331	344	BY SIMILARITY.
			FT	DISULFID	
			346	359	BY SIMILARITY.
			FT	DISULFID	
			365	376	BY SIMILARITY.
			FT	DISULFID	
			372	385	BY SIMILARITY.
			FT	DISULFID	
			387	400	BY SIMILARITY.

RESULT 7
MTN2_HUMAN

STANDARD; PRT: 956 AA.
 12 Q00339; Q9NSZ1;
 AC 000339; Q9NSZ1;
 DT 30-MAY-2000 (Rel. 3.9, Created)
 DT 16-OCT-2001 (Rel. 4.0, Last sequence update)
 DT 28-FEB-2003 (Rel. 4.1, Last annotation update)
 DE Matrin-2 precursor.
 GN MTN2.
 OS Homo sapiens (Human).
 OC Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N A. (ISOFORMS LONG AND SHORT).
 RX PubMed1124542;
 RA Muratoglu S., Krysan K., Balazs M., Sheng H., Zakanay R., Modis L., Kiss I., Deak F.;
 RA "Primary structure of human matrin-2, chromosome location of the MTN2 gene and conservation of an AT-AC intron in matrin genes.",
 RA Cytogenet. Cell Genet. 90:323-327(2000).
 RN 12
 RP SEQUENCE OF 644-956 FROM N A.
 RX MEDLINE:9728863; PubMed=9083061;
 RA Deak F., Piecha D., Bachrati C., Paulsson M., Kiss I.;
 RT "Primary structure and expression of matrin-2, the closest relative of cartilage matrix protein within the von Willebrand factor type A-

Accession	Description	Sequence	Length	Organism	Protein ID	Score
224 COHSDDDTADGPECSCHPQYKHEHTDGRSCLEREDDTLEVTESENNTSVVDGKRVKRJLM		283				
1662 --TCINVGNYTICPPDYMQNGNNCMMDRRSLCVRNYYADNQTCDGELLFMNTKCKM		1718				
284 ETCAVANING-GCBPTCK-		364				
1719 CCCSYNIGRAWNKPCEQPISTDEFATLGCSQRPGFVID:YTGLPV2IDECRE1PGYCE		1778				
305 -----HOSCPVGFQLQDGKTKMDIBEQ-----		328				
1779 NGVCIAMYGSPECEPQGPFYNDKLIVELDIDEGQNPUCORNAECACTTAGSSYRCDCRKS		1938				
329 ---TRNGCCDH-----		371				
1639 YRFSTSGOQNDRNEQELPNICSHGQCDDTVGSPYCCTIGEFTKNAQDMQCDINECERD		898				
372 RTCDHSCINHPOTFAACNRGTYLGYFTHCGDNECSNNGG-QQ-NCVNTGSSYECGC		429				
1699 AGCGNTCTRNTIGSFNCRCHGFFLHSRNNDIYDVECATAGNGLCRNGQINTGFSFQGCC		956				
430 HFGYKLAHNKDCDVEVKGL--PTSVPSPVSLHGKGSGGGDFGLRCHSGHISDVTTI		487				
1959 NCGYEVARPDGRTVIDNECLLEFKCAPO----CONLDGSYRC-----		2012				
488 RTSTVTFKLINE-----GKCSLKNQAE-----LFEPEGLR-PALPEKHSVSKESFRYVN		531				
2013 -----D1DECVEPEPICALGTC-----NTEGSFKCJPDGFSLSSTGRRQCDLNSYCAK		2664				
532 JTCSGGKOVPGNAGBRESPTRKEMETIVEELETNOKEUTASCDL-----CIVKRTKX		584				
2065 F-----EGKCKSSPK-----SRNEHSKOECCCCALKGEWGMGDPCELCPTTEP-----2103						
585 LRKAIRTLRKAVERQFHLQL-----GMMNDVAKKPRTSBRQAQS-----626						
2104 -----DEARQICPYCSGII1VGPDSAVMDCEKDPCDVKHKSGQCLNTDGSY		2149				
627 ---CGVGGQGHAEQCV---SCRAGTYDGAERCLCPNGTFTQEBCGOMTCEP-SPRGNS		680				
2150 RCBCPFVGTILEGNBCVDTECSVGN-----PGCNGTQCRVIGGEC-TCEBGFPP		2198				
681 GAIKTPPEAWRMSEBCGG--LICQDPEYSANDGAPCQLCAIGTQFDEAGRIFSC--FPCGEG		734				
2199 GPMTKCE-DINECAQNPLCAFCRVCNTYGSYECK-CPTGIVLRE-DRRMCKDDEDECEBG		2254				
735 LATRKHQGATSFQDCETRYQ---CSPGHFYNTTTHCRCPVTHYQPEFGKNCVSCEP		789				
2255 ---KHDCAEKQMECKNLAGMYICUICGQYQRRPDPGEGCV-----DNECQTFRG		2160				
790 NTITDFGDSNTITQCKKONRCCGGELGDFGYIESPNYPGNXPANTEF--WTINPPKERR		847				
2301 -----ICENGFLCLNTRGFSY-----CBNDQFTAS-----2325						
848 LIVVPEIPLIEDCGD-----YLVMRKTSSSSNVITYETCQTYER-----888						
2326 -----PTQDCELDNREGYCFTEVQLQNMCOIGSSRNRPVTKSECCCDGGRGWGPHE		2376				
889 --PIAFTSSSKLWIQFKSNEGNPSARGFQVPPYYDEDYQELIEDIVRDG		936				
2377 ICQFQGTVAFKKL-----PHGRGFMNTGADIE--CKV1HDCVRNG		2456				

۹۱

CHAIN	21	LATENT TRANSFORMING GROWTH FACTOR BETA BINDING PROTEIN, ISOFORM 1S.	POTENTIAL	POTENTIAL
FTT	300	340	EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
REPEAT	348	412	REPEAT A.	(POTENTIAL).
DOMAIN	546	587	EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
FTT	588	629	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
DOMAIN	630	670	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
FTT	671	710	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
DOMAIN	711	751	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
FTT	752	792	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
DOMAIN	793	833	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
FTT	834	874	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
DOMAIN	875	916	EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
FTT	917	958	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
DOMAIN	959	1001	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).	(POTENTIAL).
REPEAT B.	1017	1084	REPEAT B.	
DOMAIN	1097	1139	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).	
REPEAT C.	1190	1262	REPEAT C.	
DOMAIN	1140	1180	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).	
DOMAIN	1294	1334	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).	
DOMAIN	1335	1379	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).	
SITE	947	949	CELL ATTACHMENT SITE (POTENTIAL).	
DISULFID	304	315	BY SIMILARITY.	
DISULFID	310	324	BY SIMILARITY.	
DISULFID	326	339	BY SIMILARITY.	
DISULFID	350	562	BY SIMILARITY.	
DISULFID	557	571	BY SIMILARITY.	
DISULFID	573	586	BY SIMILARITY.	
DISULFID	592	604	BY SIMILARITY.	
DISULFID	599	613	BY SIMILARITY.	
DISULFID	615	628	BY SIMILARITY.	
DISULFID	634	645	BY SIMILARITY.	
DISULFID	640	654	BY SIMILARITY.	
DISULFID	657	669	BY SIMILARITY.	
DISULFID	675	686	BY SIMILARITY.	
DISULFID	681	695	BY SIMILARITY.	
DISULFID	698	709	BY SIMILARITY.	
DISULFID	715	726	BY SIMILARITY.	
DISULFID	721	735	BY SIMILARITY.	
DISULFID	737	750	BY SIMILARITY.	
DISULFID	756	767	BY SIMILARITY.	
DISULFID	762	776	BY SIMILARITY.	
DISULFID	778	791	BY SIMILARITY.	
DISULFID	797	808	BY SIMILARITY.	
DISULFID	803	817	BY SIMILARITY.	
DISULFID	819	832	BY SIMILARITY.	
DISULFID	838	850	BY SIMILARITY.	
DISULFID	845	859	BY SIMILARITY.	
DISULFID	861	873	BY SIMILARITY.	
DISULFID	879	891	BY SIMILARITY.	
DISULFID	895	906	BY SIMILARITY.	
DISULFID	902	915	BY SIMILARITY.	
DISULFID	921	933	BY SIMILARITY.	
DISULFID	927	942	BY SIMILARITY.	
D-SULFID	944	957	BY SIMILARITY.	
DISULFID	963	975	BY SIMILARITY.	
DISULFID	970	984	BY SIMILARITY.	
DISULFID	986	1000	BY SIMILARITY.	
DISULFID	1101	1114	BY SIMILARITY.	
DISULFID	1109	1123	BY SIMILARITY.	
DISULFID	1124	1138	BY SIMILARITY.	
DISULFID	1144	1155	BY SIMILARITY.	
DISULFID	1150	1164	BY SIMILARITY.	
DISULFID	1166	1179	BY SIMILARITY.	
DISULFID	1298	1309	BY SIMILARITY.	
DISULFID	1304	1318	BY SIMILARITY.	
DISULFID	1320	1333	BY SIMILARITY.	
DISULFID	1339	1354	BY SIMILARITY.	
DISULFID	1349	1363	BY SIMILARITY.	
DISULFID	1365	1378	BY SIMILARITY.	
MOD-RES	647	647	HYDROXYLATION.	
CARBONYL	810	21	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBOHYD	52	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBONYD	98	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBONYD	294	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBONYD	870	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBONYD	923	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CARBONYD	1039	N-LINKED (GLCNAC).	(POTENTIAL).
FT	CONFLICT	365	FTID-CAR	00184.
FT	CONFLICT	504	Y -> H	(IN REF. 2).
FT	CONFLICT	1334	A -> T	(IN REF. 2).
SC	SEQUENCE	1394	AA;	152791 MW; DFCASLA40B4C7C; CRC64;
Query Match		11.-%	Score 624.5;	DB 1;
Best Local Similarity		24.9%	Pred. No. 6-34;	Gaps 48;
Matches 24C		Conervative 94;	Mismatches 334;	
Qy	45	DVDECAQQGLDDCHADALCQNPTSYKSCCKGYQ-	GEGRCED3DCCGNELNGGCVHD-	101;
Dy	546	EINECTYNPDEIGA-GHCTNLUPVRYCIVCPTCYEGYRFSEQRKVCDK-	TQVCHLCSQGR	603
Qy	102	CUNIPGNYRRTCFTGEMJAHDHGNCLVDRENNNGCCTVNNGSYE-CCKEFFF-	160	
Dy	604	CENTEGSPLCZPAGPHASEBTGNC-DVDCURPDYGEHCWNTGAFREYCDSGYRM	663	
Qy	161	SDNQHTCHRSEEGLSGMNKDHGCSHICKAEPGSVACECRGFELAKNQRDCLTCNHG	220	
Dy	664	TQR----GRCEDIDELNPSTCPDECVNSP-GSYQC-----VPCTEG	7C;	
Qy	221	NGCOHSSCJDADGPESCH:POYKMETGCSRJERDTW:LEVTESENITSVTDGDVKVRR	26C	
Dy	702	FRGANGQCLDVE-----CLE-----	717	
Qy	28:	L-METCAVNGCQRTKCT:7GVHSSCPVGFTQLDGKTKD13CCTRNNGCCEHFCKN	340	
Dy	718	--PNVUC--NG---DGSNLLEGSYMWSCHKGVTRTD:K43CQDRD15CQQGNLCUNGQCKN	769	
Qy	341	IVGSGOOGCKKGFKL11TEKSCQDVDEBCSLDRTCPH--SCINHPC7FACACNRGTYLGF-	398	
Dy	770	TGSSFRTCGGQYQJSAAKQ3CQEDICEQHRHICAE3QCNTEGFSQVCDQGRASSG	829	
Qy	393	THCGDTNECS1NNGGC5-Q-VGNTVSSYECQCHPGYKLNK-----KCDEVKGLLPTS	452	
Dy	830	DH-BDINECLEKSVQRGDCNTASYDCTCPDGQLDNKTCCQDNEC-EHPGSL---	584	
Qy	452	VSPRVSHCGKGS--3GDGF-LRGSGTHLSSDTVIRTSVTFK:NEGCKSLXNAEJF	508	
Dy	885	--CGPO3ECLNTEGSHCVCQGFSISADRT-----CSDI	918	
Qy	529	PEGLRPALPEKH--SSVKEKSFYRNLTCSSGKQV-----GAPCRP-	547	
Dy	976	ENVEGSPFLCVCADENYSY2:GCCRSRTSTDYD2DOPKEEKYYLNDAISLCNVN	1035	
Qy	596	--VHREQPHIQLS--GMNLDVAKPRPRSERQNESGVGCGH-----AEN-	636	
Dy	1036	LAPVNTROECCTSQAGNDCE:FCPPVLTGAETTEMCPXKGKGFPVAESESSEAGENY	1C95	
Qy	549	STPKEXFLTV-----EFELETKOEVTAASDLSCIVKPFKJRKIAIRLRAK-----595		
Dy	1156	VNTEGSYNCFCFHPLDASEKRCIRFAESNEQIEETDVYDOLCWHEHLSDYEVCSPRP	1215	
Qy	682	ALKT-----PEAMNSCEGGJCPGEYSAQDFAPCOLALGTFQPEAGRSTCF-	729	
Dy	1036	KODABCLLFGQEICKNGFCLNTRPGYECYCKGTYDPVKQCFEDMDECQDSSIDGQC	1155	
Qy	563	QNEEGQMTC-----EPCPRPGNS-----G	681	

Db 1272 QYTPPEADPYFIDRFLNSFEELQAE-EC--GILNGCENGROVRVQEGYT:CCCDGYYHLST 1328
 Qy 780 GKKNCVSPGNTTDFGSTMN-ITQCKNRGGELGDF----TGYIES--PGNY 829
 Db 1329 AKMTCFD----VNECDELNRMMSLCKNAKCLNTDGSYKCLCPGIVPSDRNPVCTPNT 1382
 Qy 830 PANTE 834
 Db 1384 ALNLE 1388

RESULT¹⁰
LTBL_HUMAN STANDARD;
 AC Q14766; DT 28-FEB-2003 (Rel. 41, Created); DT 15-SEP-2003 (Rel. 42, Last annotation update); DE Latent transforming growth factor beta binding protein, isoform L1 precursor (LTBP-1) (transforming growth factor beta-1 binding protein DE 1) (TGF-beta1-BP-1).
 GN LTBP1.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; CC Mammalia; Butherzia; Primates; Catarrhini; Hominidae; Homo. OX NCBI TAXID=9606;
 RN [1] SEQUENCE OF 1-346 FROM N.A.
 RP TISSUE-B1000;
 RC MEDLINE=96125117; PubMed=9537398;
 RX RA Olofsson A., Ichijo H., Moren A., ten Dijke P., Miyazono K.,
 RA Heldin C.-H.; RT "Efficient association of an amino-terminally extended form of human
 RT latent transforming growth factor-beta binding protein with the
 RT extracellular matrix." J. Biol. Chem. 270:31294-31297(1995).
 RL RN [2]
 RP SEQUENCE OF 347-1595 FROM N.A.
 RC TISSUE-Fibroblast, and Platelet;
 RX MEDLINE=90275601; PubMed=2350783;
 RA Kanzaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.; RT "TGF-beta 1 binding protein: a component of the large latent complex
 RT composed of the TGF-beta 1 binding protein, a component of the large latent complex
 RT of TGF-beta 1 with multiple repeat sequences." Cell 61:1051-1061(1990).
 RL RN [3]
 RP INTERACTION WITH FIBRILLIN.
 RX Pubmed=12429738;
 RA Isogai Z., Onc R.N., Ushiro S., Keene D.R., Chen Y., Mazzieri R.,
 RA Charbonneau N.L., Reinhardt D.P., Rifkin D.B., Sakai L.Y.,
 RA "Latent transforming growth factor beta-binding protein 1 interacts
 RT with fibrillin and is a microfibril-associated protein." J. Biol. Chem. 278:27520-27527(2003).
 CC -- SUBUNIT: The large latent complex of TGF-beta from platelets is
 CC composed of the TGF-beta 1 binding protein, a component of the large latent complex
 CC of the TGF-beta 1 binding protein, and a third component denoted TGF-
 CC beta-1-BP. TGF-beta-1-BP does not bind directly to active TGF-BETA-1.
 CC Birds to fibrillin.
 CC -- SUBCELLULAR LOCATION: Secreted.
 CC -- ALTERNATIVE PRODUCTS:
 CC Event=Alternative Splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoID=Q14766-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoID=P2064-1; Sequence=External;
 CC -- TISSUE SPECIFICITY: The long isoform is found in fibroblasts.
 CC -- PTM: Contains HYDROXYLATED ASPARAGINE RESIDUES (BY SIMILARITY).
 CC -- SIMILARITY: Contains 16 EGFR-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/arcanee/> or send an email to licensee@isb-sib.ch).
 EMBL: AAP96327; DR: M34052; DR: M3461163; ALT-INIT.
 DR: HSSP; DR: P08709; DR: P08709;
 DR: MIN; DR: 11C39C; DR: GC; GC: 2005578; DR: extracellular matrix; NAS.
 DR: GC; GC: 05024; DR: transforming growth factor-beta receptor ac...; NAS.
 DR: InterPro: IPR00052; DR: InterPro: IPR0052; DR: hydroxylation; NAS.
 DR: InterPro: IPR00981; DR: InterPro: IPR00981; DR: EGF-Ca.
 DR: InterPro: IPR03619; DR: InterPro: IPR03619; DR: EGF-Ca.
 DR: InterPro: IPR00212; DR: InterPro: IPR00212; DR: fibrillar-assoc.
 DR: Pfam: PF00006; DR: Pfam: PF00006; DR: EGF_15.
 DR: SMART: SM00119; DR: SMART: SM00119; DR: ASX_HYDROXYL; 13.
 DR: PROSITE: PS00019; DR: PROSITE: PS00019; DR: ASX_HYDROXYL; 13.
 DR: PROSITE: PS01185; DR: PROSITE: PS01185; DR: EGF_2; 11.
 DR: PROSITE: PS01187; DR: PROSITE: PS01187; DR: EGF_2A; 15.
 KW Glycoprotein; DR: Glycoprotein; DR: Alternative splicing.
 KW Growth factor binding; DR: Growth factor binding; DR: EGF-like domain; DR: Hydroxylation; DR: Sigma1; DR: Sigma2.
 FT SIGNAL 1 23
 FT CHAIN 24 1595
 FT DOMAIN 501 541
 FT REPEAT 549 613
 FT DOMAIN 747 788
 FT DOMAIN 789 830
 FT DOMAIN 831 871
 FT DOMAIN 872 911
 FT DOMAIN 912 952
 FT DOMAIN 953 993
 FT DOMAIN 994 1034
 FT DOMAIN 1035 1075
 FT DOMAIN 1076 1117
 FT DOMAIN 1118 1159
 FT DOMAIN 1160 1202
 FT REPEAT 1218 1285
 FT DOMAIN 1298 1340
 FT REPEAT 1391 1463
 FT DOMAIN 1341 1381
 FT DOMAIN 1495 1535
 FT SITE 1536 1580
 FT DISULFID 1548 1650
 FT DISULFID 505 516
 FT DISULFID 511 525
 FT DISULFID 527 540
 FT DISULFID 751 763
 FT DISULFID 758 772
 FT DISULFID 774 787
 FT DISULFID 793 805
 FT DISULFID 800 814
 FT DISULFID 816 823
 FT DISULFID 835 846
 FT DISULFID 841 855
 FT DISULFID 858 876
 FT DISULFID 876 897
 FT DISULFID 882 896
 FT DISULFID 899 916
 FT DISULFID 916 927
 FT DISULFID 922 936
 FT DISULFID 938 951
 FT DISULFID 957 968
 FT DISULFID 963 977
 FT DISULFID 979 992
 FT DISULFID 998 1009
 FT DISULFID 1004 1018
 FT DISULFID 1033 1033
 FT DISULFID 1039 1055

PT	DISULFID	1046	1060	BY SIMILARITY.	Db	1177 ENVEGSFLCVACDENQEYSPMTGCCRSTSSTDLDVQDQKEERKECYVNNDASLCNVN 1236
PT	DISULFID	1062	1074	BY SIMILARITY.	Qy	596 ----VHRECFFHQLS---GMNLDVAKKPRTRSEQAESCGVQGH-----AEY- 636
PT	DISULFID	1080	1092	BY SIMILARITY.	Db	1237 LAPNTVTCQECCCTSGAGENCEIFPVLGTAETMPCKGKFVPGESSEAGGENY 1296
PT	DISULFID	1086	1101	BY SIMILARITY.	Qy	637 ----QCV-----SCRAYTDGARERC1-----LCPNGCF 662
PT	DISULFID	1103	1116	BY SIMILARITY.	Db	1297 KDAECLIPFGQEICKNGFCLNTRPGYCCKRQTYDPVKLQCFLMDEODPSSCLJ3QC 1356
PT	DISULFID	1122	1134	BY SIMILARITY.	Qy	663 CNEEGQMC-----EPCPRGNS-----S 681
PT	DISULFID	1128	1143	BY SIMILARITY.	Db	1357 VATEGSYNCFTPHMLDASEKRTRPAESNEQIEETDYYQDLCLMEHLSDYEYCSRFLVG 1416
PT	DISULFID	1145	1158	BY SIMILARITY.	Qy	682 A:KT-----PEANMSECGLGCPGEYSADGFAFCQLCAGTFQPFAGRTS CF---- 729
PT	DISULFID	1164	1176	BY SIMILARITY.	Db	1417 KQTYYTECCLYGEAEMW-CC-ALC-PLKDSDDYAGLCLNPVTGRQP-YGRDALWDSE 1472
PT	DISULFID	1171	1185	BY SIMILARITY.	Qy	730 ---PCGGGLATHQGATSFQDCETRVQSPHFYNTTHRCIRCPVG-----TYQPF 779
PT	DISULFID	1187	1201	BY SIMILARITY.	Db	1473 QTPPEADYFYIQDRFLNSFEELQAE SC-GILNGENGRCVRVQEGYTDCLGYHLDT 1529
PT	DISULFID	1302	1315	BY SIMILARITY.	Qy	780 GKNNCVSCPGNITTCPDGSTN-ITOCKNRGGELGF-----TGYES - PNY -PNY 829
PT	DISULFID	1310	1324	BY SIMILARITY.	Db	1530 AKNTCFD----VNECDLNNRMSCKNNAKINTGSKCLPFGVPSDKPNYCTP.LNT 1584
PT	DISULFID	1326	1339	BY SIMILARITY.	Qy	834 EOC PANTE 834
PT	DISULFID	1345	1356	BY SIMILARITY.	Db	1555 ALNLE 1599
PT	DISULFID	1351	1365	BY SIMILARITY.	RESJCT 11	RESJCT 11
PT	DISULFID	1367	1380	BY SIMILARITY.	QES_XCJSE	QES_XCJSE
PT	DISULFID	1499	1510	BY SIMILARITY.	AC	2-IDS_XCJSE STANDARD: PRT: 1389 AA.
PT	D.SULFID	1505	1514	BY SIMILARITY.	AC	Q8CG1B_Q8BNW7_Q8CIRC;
PT	DISULFID	1521	1534	BY SIMILARITY.	DT	15-SEP-2003 (Re: 42, Created;
PT	DISULFID	1540	1555	BY SIMILARITY.	DT	15-SEP-2003 (Re: 42, Last sequence update)
PT	DISULFID	1550	1564	BY SIMILARITY.	DT	15-SEP-2003 (Re: 42, Last annotation update)
PT	DISULFID	1566	1579	BY SIMILARITY.	DE	Latent transforming growth factor beta binding protein precursor (LTBP-1); Transforming growth factor beta-binding protein
PT	MOD-RES	848	848	HYDROXYLATION (BY SIMILARITY).	DE	(TGF-beta-1-3p1).
PT	CARBONYD	1011	1011	N-LINKED (GLCNAC. .) (POTENTIAL).	GN	LTBP1
PT	CARBONYD	495	495	N-LINKED (GLCNAC. .) (POTENTIAL).	CS	Mus musculus (Mouse)
PT	CARBONYD	1071	1071	N-LINKED (GLCNAC. .) (POTENTIAL).	OC	Exaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
PT	CARBONYD	1124	1124	N-LINKED (GLCNAC. .) (POTENTIAL).	OC	Xenialia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
PT	CARBONYD	1240	1240	N-LINKED (GLCNAC. .) (POTENTIAL).	RN	[1].
SEQ	SEQUENCE	1595	AA;	173229 MN;	RN	PubMed-ID:12090
Qy	Query Match	11:1:	Score 624.5; DB 1; Length: 1595;		RZ	SEQUENCE FROM N.A.
	Best Local Similarity	24.9%	Pred. No. 7e-34;		RZ	PubMed-ID:12711388;
	Matches	240;	Conservative	94; Mismatches 334; Indels 297; Gaps 48;	RZ	Kugler I., Obata H., Gualandris A., Cowin P., Riekin D.B.,
Qy	45 DVDECAQ-DDCHADALCQNTPTSYKCSCKPGYQ--GEGRCEDIDECKGNELNGCVCHD-	101			RZ	"Molecular cloning of the mouse Ltbp-1 gene reveals tissue specific expression of alternatively spliced forms."
Db	747 E.NECTYNPDICOA-GHCINLPLVRYTCIYBEYRFSEQQRCVCD:DEC-TQVQH:CSQR	304			RZ	Gene 308:31-41 (2003). [2]
Qy	102 C*NIQPNYRCTCFDFGMLAHGHNCLDVECLEENNGCOH:CNVNMGSYEC-CCKEGFFL	160			RZ	SEQUENCE FROM N.A.
Db	805 CENTEGSFLCIPAGMASEEGTNCDVDECLRDPVGEGEHCNTVGAPECYCPSGYRN	364			RZ	Strain: C57BL/6J; Tissue=Liver, and Vein;
Qy	161 SDNQHTCITHREEGSCLSMNHDGSHICKEAPRSVACECPFPFLAKNORDCLTCHG	220			RZ	Weiskirchen R., Moser M., Guenther K., Weiskirchen S., Gressner A.M.,
Db	865 TQR----GRCEID:DECINPSTCPDCECVNSP-GSYC-----VPCTES	902			RZ	"The murine latent transforming growth factor-beta binding protein, Ltbp-1, is alternative, transiently spiced, and maps to a region syntenic to human chromosome 2p21.22."
Qy	221 NGCQHSDDTAGPESCHPKQYKMTDGRSCLEREDTVLTVTESNTISVYDGDKRVKSR	28C			RZ	Gene 308:43-52(2003). [3]
Db	903 FQWNGQCLDVE-----CLE-----	91:8			RZ	SEQUENCE OF 788-1389 FRM N.A.
Qy	281 LLMETCAVNGGCDRTCKDTSVHCHSCPVGFTLOLDGKTKDQEDCQTRNGGCDPHCKN	340			RZ	Strain: C57BL/6J; Tissue=Aorta, Liver, and Vein;
Db	919 --PWNCA--NG---DSNLESSYMC5CHKGTRDHKRDQEDCQGQSLCYNGQCKN	970			RZ	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Qy	341 IVSFDGCKKGFKLTLDEKSQDVBCSLDRTCDH-SCINHPGTFAACMRGTYGF-	398			RZ	Okada I., Osato N., Saito R., Nakao T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Db	971 TEGSFRCTGQGSAQDQECDIDCQHRLCAHGCCRNTEGSFQCVQDQGRASGLG	1030			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Qy	399 THGDTNBCSINNGGCOO-VCINTVGSYBQCHPQGYKLHWNK----KQCVEVKSLPLTS	452			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Db	1031 DHEDINECLEDKSVQRGDCINTRGSYDCTCDNKTCDQINEC-EHPLG-----1C85				RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Qy	453 VSPRVSLHGKSG---GGGCF-LRCHSGHILSSDVTIRTSVFKLNEGRSLXNAELF	508			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Db	1086 -----CGPQECIANTEGSFTCVCQGPFSISADRT-----CEDI	11:13			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Qy	509 PEGLRPALPEKH---SSVYKESFRYVNLTCSKGQVP-----GAPGRP--	547			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Db	1120 DECVNNTCDSHGFCNDTAGSFRCL---CYQFOAQARQDGQGCVDNCEC:LSGVGEAFC	1176			RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Qy	548 STPKEMPFTY----EPELEINQKEVNTASSCDLSCIVKTRKLKAIRLTKA-----595				RZ	Okada I., Osato N., Saito R., Suzuki H., Yamamoto T., Noguchi C., Gojobori T., Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beise K.W., Brake C.A., Bradt D., Bruylants C., Corbin J.E., Cousins S., Dalla E., Draganic T.A., Fletcher C.F., Forrest A., Frazer K.S.,

EMBL; AF346463;	AAN77251.1;	JOINED;
EMBL; AF346464;	AAN77251.1;	JOINED;
EMBL; AY143161;	AAN78831.1;	ALT_SEQ.
EMBL; AKC5038C;	BAC3422.1;	-
EMBL; AKC80224;	BAC7808.1;	-
YGBL; MGL1_109251;	LDEP;	-
InterPro; IPRO0152;	Asx hydroxy;	-
InterPro; IPRO0188I;	EGF CA;	-
SMART; SM03179;	EGF CA;	16.
PROSITE; PS00100;	ASX HYDROXY;	13.
PROSITE; PS00022;	EGF 1;	1.
PROSITE; PS01186;	EGF 2;	1.
PROSITE; PS01187;	EGF-CA;	15.
Growth factor binding; Repeat; EGF-1;		
Hydroxylation; Sigma-1; Glycoprotein;		
Allerterative splicing;		
SIGNAL;	1 19	POTENTI-
CHAIN;	20 139	LATENT;
DOMAIN;	295 335	BINDING;
REPEAT DOMAIN;	343 408	EGF-LIK
542 583	(PCTEN)	
DCM1N;	584 625	REPEAT
DCMAIN;	626 666	EGF-LIK
DCMAIN;	667 706	(POTENT)
DOMAIN;	707 747	EGF-LIK
DOMAIN;	749 798	(PCTEN)
DOMAIN;	769 829	EGF-LIK
DOMAIN;	830 870	EGF-LIK
DOMAIN;	871 912	(POTENT)
DCM1N;	913 954	EGF-LIK
DCM1N;	955 997	(PCTEN)
REPEAT DOMAIN;	1012 1079	EGF-LIK
DCMAIN;	1092 1134	REPEAT
DOMAIN;	1135 1175	EGF-LIK
DISUFID;	299 313	(POTENT)
DISUFID;	305 319	POTENTI-
DISUFID;	321 334	POTENTI-
DISUFID;	546 558	POTENTI-
DISUFID;	553 567	POTENTI-
DISUFID;	569 582	POTENTI-
DISUFID;	588 600	POTENTI-
DISUFID;	595 609	POTENTI-
DISUFID;	611 624	POTENTI-
DISUFID;	630 641	POTENTI-
DISUFID;	636 645	POTENTI-
DISUFID;	653 665	POTENTI-
DISUFID;	661 673	POTENTI-
DISUFID;	682 694	POTENTI-
D-SU-FID;	677 691	POTENTI-

PT	DISULFID	694	705	POTENTIAL.	Db	1230 C-ALCPKMD--SDDYA---QLCNI-----PVTGRRPVGDA-VDFSEQYGPRTDPPYF-QDR
PT	DISULFID	711	722	POTENTIAL.	Qy	741 GATSFQDETRIQCSPGFHYNTTHJCRCPIG----TYQPEFGKNCVSCCGNTT 793
PT	DISULFID	717	731	POTENTIAL.	Db	1281 FLSNFEELQAE-EC-GI1NGCENGRCYRHGEYTCDFSYHLMAKTCYDV--NECS 1335
PT	DISULFID	733	746	POTENTIAL.	Qy	794 DFDGSTNTQCNRRCGELGD-----TGYLES----PEAWNME 834
PT	DISULFID	752	763	POTENTIAL.	Db	1336 ELNNRMSL-CRKAICNTEGSYKLCLPGYIPSDFKNYCTPLNSAIND 1385
PT	DISULFID	758	772	POTENTIAL.		
PT	DISULFID	774	787	POTENTIAL.		
PT	DISULFID	793	804	POTENTIAL.		
PT	DISULFID	799	813	POTENTIAL.		
PT	DISULFID	815	828	POTENTIAL.		
PT	DISULFID	834	846	POTENTIAL.		
PT	DISULFID	841	855	POTENTIAL.		
PT	DISULFID	857	869	POTENTIAL.		
PT	DISULFID	875	887	POTENTIAL.		
RESULT 12						
LTBP1_MOUSE STANDARD: PRT: i713 AA.						
Query Match		10.7%	Score 601; DB 1; Length 1389;	AC Q8C919; Q88349; Q9BAX7; Q837FS; Q8CIR0;		
Best Local Similarity		24.2%	Pred. No. 2,2e-12;	DT 15-SEP-2003 (Rel. 42, Created)		
Matches		230;	Mismatches 305; Indels 310; Gaps 52;	DT 15-SEP-2003 (Rel. 42, Last sequence update)		
Qy		45	DVDECAQQLDCHADALCQNTPHYSYKCSCKRGQG--EGRQEDIDEPCGN--LNGGV 99	DT 15-SEP-2003 (Rel. 42, Last annotation update)		
Db		584	DIVECAQVRLHC-SQRGCENTEGSFLCUPGMAESEGKCIDVDECLRDMCPBDGRCI 642	DE Latent transforming growth factor beta-binding protein, isoform 1L precursor (LTBP1) (Transforming growth factor beta-binding protein)		
Qy		100 H	-DCLNIPGNYRCT-CFDGFMLAH 121-	DE (TGF-beta-BP1).		
Db		643	NTAGAFRECYCDGYSRGSRRGYCEJDIECLKPSLTCPEQCNTPGSQCVFTEGR-- 699	DN LTBP1_TaxID:3090;		
Qy		122	DGHN--CLDVDCELENNGGCOHTCVCNVMGSYECCKEGFFLSDQHCTC--THRSEEGLSC 277-	RN [1];		
Db		700	-GWNQGCLDVDCELQPKVCTNGSCTNLEGSTMCSHRGSYSPTRDHHCCDQDECCQ3NL 758	SEQUENCE FROM N.A.		
Qy		178	MNDHGSHCIIKEAPRGSSVACECRGPGLF--AKNQRCDCITLNHGNGGCCII-SCD7ADG 224	SEQUENCE FROM N.A.		
Db		759	MNGQ-----CRNTD-GSFRCTGGQGQLSAAKDQCCEDIDECEH-HHLCSHGOCNTEGS 810	STRAIN: 29/385;		
Qy		235	PEGSCHPPOKSHITGCRSLERDQTLETVTEENTSVWVGDKRVKRLLMECTAVNNGCC 294	RA PubMed:1271388;		
Db		911	FQCVCNQGYASVLGDHC--ED--INE-----CLESSVSGQG-- 844	RA Weisskirchen R., Moser M., Quenther K., Weiskirchen S., Gressner A.M.; Weisskirchen R., Moser M., Quenther K., Weiskirchen S., Gressner A.M.; LTBP1 is alternatively spliced, and maps to a region syntenic to human chromosome 2p11-22.;		
Qy		295	RICKDTSTGVHCSCPVGFTLQDGKTKCDIDEQCTRNNGCCHF--SKNIYGSFDGCKKG 352	RA Gene 306:3-41 (2003);		
Db		845	-DCINTASSYDPCTCPDGFLQN-DNKGCDINNE-AOPGLCGSRGECCINTQSFHCVCEQ 9C1	[2];		
Qy		353	EKLIDDKSCDQYDECSLDRTCD-HS-CINHNGTFAACRGYTL-YGFTHGCDNECSI 405	RA SEQUENCE OF 1112-1713 FROM N.A.		
Db		902	FS-SADGRTCD-DECYNTVCDSHGFCDNTAGSFRCLQYQFOAQPDQGCVDNECEL 961	RA STRAIN: G57BL/6C;		
Qy		410	NNGGC-QQVCMVNTGVSYECQC--HPGY----KJHNKKKDC---V 443	RC TISSUE: Lorta, Liver, and Vein;		
Db		962	LSGVCGEAFCENVEGGSFLCVADENOEYSPMTGQCRSRVTEPSGVDQRPREKKEYYNL 1021	RC		
Qy		444	EVKGLLPTSVSPRVLH--CGKGGGDDCFLRCHSGIHLSSDVTIRTSVFKLNEGKCS 50;	RA PubMed:12466851;		
Db		1022	NDASLCNDVNLAPNVTQBECCCTSGAGG-- 1049	RA Okazaki Y., Furuno M., Kasukawa T., Adachi T., Bono H., Kondo S., Maekawa I., Osato N., Saite R., Suzuki R., Yamazaki T., Kiyosawa H., Niwa H., Nakao A., Nagami A., Schonbach C., Gotohori T., Baldarelli R., Hii D.P., Bult C., Hune D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Bataivo S., Beisel K.W., Blake C.A., Bradt D., Brusick V., Chothia C., Corbeni L.E., Cousins S., Dali E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Garibaldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gusinovich S., Hirokawa N., Jackson T.J., Jarvis E.D., Sandelin A., Schneider C., Semple C.A., Seto M., Shishada K., Konagai A., Kurochikai H., Kawano T., Kuroki R.M., King B.L., Lenhardt B., Lyons P.A., Magioli D.R., Mattais L., Marchionni L., McKenzie J., Mikl H., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Weils C., Wilming L.G., Wyrshaw Boris A., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Komio H., Nakamura M., Sakazume N., Sato K., Shiraki T., Wakai K., Wakai J., Aizawa K., Itoh M., Kagaawa T., Hara A., Hasegawa W., Imai K., Ishii Y., Itou M., Kagaawa A., Miyazaki A., Sakai K., Sasaki D., Shihara K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y., RA		
Qy		599	EQFHQLQSMNLDVAKRPRPTSERQAECCGVGOGHAENQCS-----CRAGTYDGA 650	RA		
Db		1134	QVPGAPGRPSTPKEMFTVFEFELETNQEVTASCDLSCIVKTEKLRKAIRTLKAVIR 598	RA		
Qy		1109	CNTQPG----YECYCKQSTYDPEVKLOC----- 1133	RA		
Db		502	LKNRBLFP-----EGLRPALEPKHSSVYKESFR-----YVNLTCSGGK 538	RA		
Qy		1050	-DNEIFPPVPGQTAEFFMCPCRGKLYPAGESSYDGTGENYKDABCLLGEETCKNGY 1108	RA		
Qy		539	QVPGAPGRPSTPKEMFTVFEFELETNQEVTASCDLSCIVKTEKLRKAIRTLKAVIR 598	RA		
Db		1134	-----FMDMEQCDP-----NSC-----LNGQCNTEGSYNCFTPHYLDA 1176	RA		
Qy		651	RERCILCPNGTFQEEGOMCCEP-----PRPGNSGALKT-----PEAWNME 693	RA		
Db		1171	EKRCVQPTTSENEQ-EETDYYQDUCWHELUSEVCSRPLUVGKOTTYBCCCLYGEAWGM-Q 1229	RA		
Qy		694	CGGLCQPESEYSSADGFAPCOLCALGTFQEAGRTSCF-----PGGGGLATQ:Q 740	RA		

"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RT RT Nature 422:563-573 (2002).;

RL RL SUBUNIT: The large latent complex of TGF-beta: from platelets is composed of the TGF-beta1 monomer noncovalently associated with a disulfide-bonded complex of a dimer of the N-terminal propeptide of the TGF-beta1 precursor and a third component denoted TGF-beta1-BP. TGF-beta1-BP does not bind directly to active TGF-BETA1.

CC CC SUBCELLULAR LOCATION: Secreted (By similarity).;

CC CC ALTERNATIVE PRODUCTS: Event/Alternative splicing; Named isoforms=2; Name=L; IsoID=Q8CG19-1; Sequence=Displayed;

CC CC PTM: Contains hydroxylated asparagine residues (By similarity).;

CC CC SIMILARITY: Contains 16 EGF-like domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR SMART; SM00179; EGF CA; 16.5%; PROSITE; PS00310; Asx Hydroxyl; 13. DR PROSITE; PS00222; EGF 1; 2. DR PROSITE; PS00222; EGF 2; 1. DR PROSITE; PS01186; EGF 1; 1. DR PROSITE; PS01187; EGF CA; 15. KW Hydroxylation: Signal; Repeat: EGF-like domain; KW Alternative splicing; Glycoprotein; KW Latent-transforming growth factor beta; POTENTIAL.

FT SIGNAL 1 23 FT CHAIN 24 1713 FT DOMAIN 619 659 FT REPEAT 667 732 FT DOMAIN 866 907 FT DOMAIN 908 943 FT DOMAIN 950 990 FT DOMAIN 991 1030 FT DOMAIN 1032 1071 FT DOMAIN 1072 1112 FT DOMAIN 1113 1153 FT DOMAIN 1154 1194 FT DOMAIN 1195 1236 FT DOMAIN 1237 1278 FT DOMAIN 1279 1321 FT REPEAT 1336 1403 FT DOMAIN 1416 1458 FT DOMAIN 1459 1499 FT REPEAT 1509 158. FT DOMAIN 1613 1653 FT DISULFID 623 634 FT DISULFID 629 643 FT DISULFID 645 658 FT DISULFID 670 882 FT DISULFID 877 891 FT DISULFID 893 906 FT DISULFID 912 924 FT DISULFID 919 933 FT DISULFID 935 948 FT DISULFID 954 965 FT DISULFID 960 974 FT DISULFID 977 989 FT DISULFID 995 1006 FT DISULFID 1001 1015 FT DISULFID 1018 1029 FT DISULFID 1035 1046 FT DISULFID 1041 1055 FT DISULFID 1057 1070 FT DISULFID 1076 1087 FT DISULFID 1091 1096 FT DISULFID 1101 1111 FT DISULFID 1117 1128 FT DISULFID 1123 1137 FT DISULFID 1139 1152 FT DISULFID 1158 1170 FT DISULFID 1165 1179 FT DISULFID 1181 1193 FT DISULFID 1199 1211 FT DISULFID 1205 1220 FT DISULFID 1222 1235 FT DISULFID 124.1 1253 FT DISULFID 1247 1262 FT DISULFID 1264 1277 FT DISULFID 1283 1295 FT DISULFID 1290 1304 FT DISULFID 1316 1320 FT DISULFID 1420 1433

Query Match 12.7%; Best Local Similarity 24.2%; Pred. No. 2,8e-32; Mismatches 305; Indels 310; Gaps 52; Matches 235; Conservative 105; Vismatches 305; SMART: SM00179; EGF CA; EGF-like; InterPro: IPR006209; EGF-like; InterPro: IPR002212; Fibril-assoc.; InterPro: IPR006210; IEGF; InterPro: IPR0008; EGF; 16. Pfam: PF00683; TB; 4. SMART: SM00181; EGF; 18.

PT	REPEAT DOMAIN	1517	INTERNAL REPEAT 4.	QY	95	N-----GGC-----VHDLNIPGNYRCT-CFCGMLAH 121
PT	DISULFID	1612	EGF-LIKE 17. CALCIUM-BINDING (POTENTIAL).	Db	966	STPAGAFRCYEDSGYMSRGHED-DECUTPSTCDEEYCNSPOSSYCYCFTEGFR-- 1022
PT	DISULFID	1653	1697 EGF-LIKE 18. CALCIUM-BINDING (POTENTIAL).	Db	122	2GHN--CLDVECLENGGCOHTCVNMGSYECCKEGFFLSNDHTC-IHRSEGLSC 277
PT	DISULFID	185	195 BY SIMILARITY.	QY	1223	-GWNGQCLADVECLQPXKTCNGSCTNLEGSTYMSCSKGVDPTCHRHQD-DECQGNLIC 1081
PT	DISULFID	189	201 BY SIMILARITY.	Db	1062	MNGQ----CRNTD-GSFRCGCGOYQQLSAKQDFDICEFHHR-LCSHGQCRNTEGS 2133
PT	DISULFID	203	212 BY SIMILARITY.	QY	178	XNKDHGSCHICKEAPROSVACERGPFEL-AKQEDCCTTNGNGCCQH-SCDDTAQG 234
PT	DISULFID	395	405 BY SIMILARITY.	Db	225	PESCHHQYKHTDGSCLEREDTVLEVTEANTSYSVGDKVRKRLJWETCAVNGGCD 294
PT	DISULFID	399	411 BY SIMILARITY.	QY	235	FQCLCNQGYASVLDHQ---ED---INE-----CLEDSSVCGG-- 1167
PT	DISULFID	413	422 BY SIMILARITY.	Db	1134	FQCLCNQGYASVLDHQ---ED---INE-----CLEDSSVCGG-- 1167
PT	DISULFID	622	623 BY SIMILARITY.	QY	295	RTKQTKTSGVHCGSPVGFITLGDKTKDIBCOTRNGCDEH-SKVNVSFDCGCKKG 352
PT	DISULFID	628	642 BY SIMILARITY.	Db	1158	D-INNTASGYDCTPQJGLN-DNGQDINBC-AOPGLCAPHGELNTQSFHVCQG 1224
PT	DISULFID	644	657 BY SIMILARITY.	QY	353	FKLTDEKSCOPYCECSJLRTCD-HS-CINHPGTFAACNRGTYL-YGFTPHGCFNECSI 409
PT	DISULFID	669	881 BY SIMILARITY.	Db	1225	FISIADGRTCDIDEVNTDVDSHGFDNTAGSFRCLCYGFQAFQDGCCVDTNEEL 1284
PT	DISULFID	876	890 BY SIMILARITY.	QY	410	NNNGC-QCVCINTVGSYECOCHPGYKL-----WNKEDC--V 443
PT	DISULFID	892	905 BY SIMILARITY.	Db	1265	LSSVCGEEFCEVNEVGSKVADENQEYSPMTGQRKKECYYNL 1344
PT	DISULFID	911	923 BY SIMILARITY.	QY	444	EYRGLLEPTSVSRSH-CGKSGGGDGFCLRCHSGIHLSSDVTTTSVT-FKLNEJKCS 501
PT	DISULFID	918	932 BY SIMILARITY.	Db	1345	NDSLCEVNLAPNVTROBCCTSAGWG----- 1372
PT	DISULFID	934	947 BY SIMILARITY.	QY	502	JKNAELFP-----EGIRPALPEKHSSVKESFR-----YVNLTGSSGK 538
PT	DISULFID	953	964 BY SIMILARITY.	Db	1373	-ENCEIFFCPVQ3TAERSSEMCPRGKGFVPAGESSYETGEVYKADEBCLLGEETCKNGY 1431
PT	DISULFID	1017	1028 BY SIMILARITY.	QY	539	OVEGARGPRSPSTPKEMF-TIVEFELETNCKEY-ASCDSCLIVRKTRTEKLRKA-RTLKAVHR 598
PT	DISULFID	1034	1045 BY SIMILARITY.	Db	1432	CNTQPGYECYKE-----GTYDYPVRLQC----- 1456
PT	DISULFID	1040	1054 BY SIMILARITY.	QY	599	EQPHLQLQGMNJDVAKKPRTSERQAEQGTMCTCP-----CRAFTYDGA 650
PT	DISULFID	1056	1069 BY SIMILARITY.	Db	1457	RECILCPNGTFONEEQMTCP-----NSC-----IDGCCNTEGSYNCFCPTHPVLDAS 1493
PT	DISULFID	1075	1086 BY SIMILARITY.	QY	651	GATSFQDCTRYQCSFPHFYNTTHRCRCPVG-----TYCPERGKNVNCSPONTT 793
PT	DISULFID	1081	1095 BY SIMILARITY.	Db	1494	EKSCVQPPESNQIEFTDVQDLCWELHSEEVVCSPRLJVGRCTTYDECCTYGEANGM-Q 1552
PT	DISULFID	1097	1110 BY SIMILARITY.	QY	694	CGG-CQPEYESQDFQCOLCALGTFOPEAERTSCF-----PGGGGLATKQ 740
PT	DISULFID	1116	1127 BY SIMILARITY.	Db	1533	C-ALCPMCD-----EDYQA-CQLCN-----PVTSRRRPYGRDA-VDFSEQYGPETDPYIQRD 1603
PT	DISULFID	1122	1136 BY SIMILARITY.	QY	741	GATSFQDCTRYQCSFPHFYNTTHRCRCPVG-----TYCPERGKNVNCSPONTT 793
PT	DISULFID	1138	1151 BY SIMILARITY.	Db	1604	FLNSFEEQAE-EC-GILNG-SENGCVRQEGY-CDCFDSYHLMWKTCYV-NECS 1658
PT	DISULFID	1157	1169 BY SIMILARITY.	QY	794	DEFGSTNTNOCKNRRCGGELDF-----TGYES-PNYPNPYPA 832
PT	DISULFID	1164	1178 BY SIMILARITY.	Db	1659	EJNNRMSL-CRNAKC-NTEGSYKCVCLPGVPSDKPNY-CTPLNT 1701
PT	DISULFID	1180	1192 BY SIMILARITY.	QY	835	RESULT FB-2 MOUSE
PT	DISULFID	1198	1210 BY SIMILARITY.	AC	AC P37869; Genu2;	
PT	DISULFID	1210	1219 BY SIMILARITY.	CT	CT 01-OCT-1994 (Rel. 30, Created)	
PT	DISULFID	1221	1234 BY SIMILARITY.	DT	DT 01-OCT-1994 (Rel. 30, Last sequence update)	
PT	DISULFID	1240	1252 BY SIMILARITY.	DT	DT 15-SEP-1993 (Rel. 42, Last annotation update)	
PT	DISULFID	1246	1261 BY SIMILARITY.	DC	DC Fibu-ir-2 precursor	
PT	DISULFID	1263	1276 BY SIMILARITY.	QY	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	DISULFID	1282	1294 BY SIMILARITY.	Db	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	DISULFID	1289	1303 BY SIMILARITY.	QY	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	DISULFID	1305	1319 BY SIMILARITY.	AC	AC P37869; Genu2;	
PT	DISULFID	1419	1432 BY SIMILARITY.	CT	CT 01-OCT-1994 (Rel. 30, Created)	
PT	DISULFID	1427	1441 BY SIMILARITY.	DT	DT 01-OCT-1994 (Rel. 30, Last sequence update)	
PT	DISULFID	1443	1456 BY SIMILARITY.	DT	DT 15-SEP-1993 (Rel. 42, Last annotation update)	
PT	DISULFID	1462	1473 BY SIMILARITY.	DC	DC Fibu-ir-2 precursor	
PT	DISULFID	1468	1482 BY SIMILARITY.	QY	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	DISULFID	1484	1497 BY SIMILARITY.	AC	AC P37869; Genu2;	
PT	DISULFID	1616	1627 BY SIMILARITY.	CT	CT 01-OCT-1994 (Rel. 30, Created)	
PT	DISULFID	1622	1636 BY SIMILARITY.	DT	DT 01-OCT-1994 (Rel. 30, Last sequence update)	
PT	DISULFID	1638	1651 BY SIMILARITY.	DT	DT 15-SEP-1993 (Rel. 42, Last annotation update)	
PT	DISULFID	1657	1672 BY SIMILARITY.	DC	DC Fibu-ir-2 precursor	
PT	DISULFID	1667	1681 BY SIMILARITY.	QY	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	DISULFID	1683	1696 BY SIMILARITY.	AC	AC P37869; Genu2;	
PT	CARBONYD	339	339 N-LINKED (GLCNAC,) (POTENTIAL).	CT	CT 01-OCT-1994 (Rel. 30, Created)	
PT	CARBONYD	370	370 N-LINKED (GLCNAC,) (POTENTIAL).	DT	DT 01-OCT-1994 (Rel. 30, Last sequence update)	
PT	CARBONYD	416	416 N-LINKED (GLCNAC,) (POTENTIAL).	DT	DT 15-SEP-1993 (Rel. 42, Last annotation update)	
PT	CARBONYD	612	612 N-LINKED (GLCNAC,) (POTENTIAL).	DC	DC Fibu-ir-2 precursor	
PT	CARBONYD	1042	1042 N-LINKED (GLCNAC,) (POTENTIAL).	QY	907 DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE 94	
PT	CARBONYD	1242	1242 N-LINKED (GLCNAC,) (POTENTIAL).	AC	AC P37869; Genu2;	
PT	CARBONYD	1357	1357 N-LINKED (GLCNAC,) (POTENTIAL).	CT	CT 01-OCT-1994 (Rel. 30, Created)	
SEQ	SEQUENCE	1712 AA;	186598 MW;	650BCEA691FD134 CRC34;	DT	DT Fibu-ir-2 precursor
QY	Query Match Best Local Similarity	23.0	Score 592; Pred. No. 1.1e-11; Mismatches 102; Indels 310; Gaps 52;	PRT: 1221 AA.		
Db	Database Matches	23.0	Conservative 102; Mismatches 305; Indels 310; Gaps 52;	PRT: 1221 AA.		

Queried sequence: FVDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE

Best Local Similarity: 23.0

Score: 592

Pred. No.: 1.1e-11

Mismatches: 102

Indels: 310

Gaps: 52

Annotations:

- Created: 01-Oct-1994 (Rel. 30)
- Last sequence update: 15-Sep-1993 (Rel. 42)
- Last annotation update: 15-Sep-1993 (Rel. 42)
- Standard: FB-2 Mouse
- Prerequisite: F3LN2
- Organism: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota
- Comments: DIDECAQGLDODCHADALCONTPTSYKCSCKPGY-QGEGRCEDIDE-----GNE

OC	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.	EMBL; AF135239; AAD34456.1;	JOINED.
OX	NCBI_TAXID=10090;	DR EMBL; AF135240; AAD34456.1;	JOINED.
RN	[1]	DR EMBL; AF135241; AAD34456.1;	JOINED.
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.	DR EMBL; AF135242; AAD34456.1;	JOINED.
RC	TISSUE-Fibroblast;	DR EMBL; AF135243; AAD34456.1;	JOINED.
RX	MEIDLINE=94064787; PubMed=6245130;	DR EMBL; AF135244; AAD34456.1;	JOINED.
RA	Pan T.-C.; Sasaki T.; Zhang R.-Z.; Faessler R.; Timpl R.; Chu M.-L.;	DR EMBL; AF135245; AAD34456.1;	JOINED.
RT	"Structure and expression of fibulin-2, a novel extracellular matrix protein with multiple EGF-like repeats and consensus motifs for calcium binding."	DR EMBL; AF135246; AAD34456.1;	JOINED.
RT	J. Cell Biol. 123:1250;	DR EMBL; AF135247; AAD34456.1;	JOINED.
RL	RN	DR EMBL; AF135248; AAD34456.1;	JOINED.
RN	[2]	DR EMBL; AF135249; AAD34456.1;	JOINED.
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPlicing.	DR EMBL; AF135250; AAD34456.1;	JOINED.
RX	MEIDLINE=99337686; PubMed=1040956;	DR EMBL; AF135251; AAD34456.1;	JOINED.
RA	Graessel S.; Sicot F.-X.; Gotta S.; Chu M.-L.;	DR PIR; A49457.	
RT	"Mouse fibulin-2 gene. Complete exon-intron organization and promoter characterization."	DR HSSP; PCC716; IAPQ.	
RT	RJ. Biochem. 263:471-477(1999).	DR SMART; SW00104; AKATC; 3.	
RJ	RN	DR PROSITE; PS00179; EGF CA; 9.	
RP	DEVELOPMENTAL STAGE.	DR PROSITE; PS00010; ASX HYDROXYL; DR InterPro; IPR000152; Asx hydroxyl.	
RX	PubMed=8505659;	DR InterPro; IPR000152; Asx hydroxyl; DR PIR; A49457.	
RA	Zhang H.-Y.; Timpl R.; Sasaki T.; Chu M.-L.; Ekblom R.;	DR InterPro; IPR00020; Anaphylatoxin.	
RT	"Fibulin-1 and fibulin-2 expression during organogenesis in the developing mouse embryo";	DR InterPro; IPR00020; Anaphylatoxin.	
RT	"Fibulin-1 and fibulin-2 expression during organogenesis in the developing mouse embryo";	DR InterPro; IPR00020; Anaphylatoxin.	
RT	Ekblom R.; Timpl R.; Chu M.-L.; Ekblom R.;	DR InterPro; IPR00020; Anaphylatoxin.	
RL	Dev. Dyn. 205:348-364(1996).	DR InterPro; IPR00020; Anaphylatoxin.	
RN	[3]	DR InterPro; IPR00020; Anaphylatoxin.	
RP	BINDING TO LAMA2.	DR InterPro; IPR00020; Anaphylatoxin.	
RX	PubMed=1002829;	DR InterPro; IPR00020; Anaphylatoxin.	
RA	Talts J.-F.; Andac Z.; Goehring W.; Brancaccio A.; Timpl R.;	DR InterPro; IPR00020; Anaphylatoxin.	
RT	"Binding of the G domains of laminin alpha1 and alpha2 chains and fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma";	DR InterPro; IPR00020; Anaphylatoxin.	
RL	Bur. J. Haematol. 67:176-184(2001).	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS CALCIUM-DEPENDENT.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! SUBUNIT: Homotrimer; disulfide-linked. Interacts with LAMA2.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! SUBCELLULAR LOCATION: Secreted; extracellular matrix.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! ALTERNATIVE PRODUCTS:	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! EVENT: Alternative splicing; Named isoforms=2;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	Comment-Additional isoforms seem to exist;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	Name=1;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	IsoId:P37889-1; Sequence=Displayed;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	Name=2; Synonyms=EGF3-less;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	IsoId:P37889-2; Sequence=VSP_00131;	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! TISSUE-SPECIFICITY: Component of both basement membranes and other connective tissues.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! DEVELOPMENTAL STAGE: The differential expression of the fibulin family contributes to the formation of molecularly distinct extracellular matrices already during early developmental stages of a large number of tissues.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! INDUCTION: Glucocorticoids suppressed mRNA expression and protein synthesis.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! SIMILARITY: Belongs to the fibulin family.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! SIMILARITY: Contains 3 anaphylatoxin-like domains.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-!! SIMILARITY: Contains 11 EGF-like domains.	DR InterPro; IPR00020; Anaphylatoxin.	
CC	-----	-----	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licence agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	-----	-----
CC	-----	-----	-----
DR	EMBL: X75285; CAA53040.1; -;	DR DISULFD; 449	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 479	BY SIMILARITY.	
DR	-----	DR DISULFD; 508	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 492	BY SIMILARITY.	
DR	-----	DR DISULFD; 509	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 511	BY SIMILARITY.	
DR	-----	DR DISULFD; 535	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 512	BY SIMILARITY.	
DR	-----	DR DISULFD; 542	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 525	BY SIMILARITY.	
DR	-----	DR DISULFD; 543	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 598	BY SIMILARITY.	
DR	-----	DR DISULFD; 610	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 619	BY SIMILARITY.	
DR	-----	DR DISULFD; 634	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 621	BY SIMILARITY.	
DR	-----	DR DISULFD; 683	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 692	BY SIMILARITY.	
DR	-----	DR DISULFD; 707	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 713	BY SIMILARITY.	
DR	-----	DR DISULFD; 726	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 735	BY SIMILARITY.	
DR	-----	DR DISULFD; 742	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 754	BY SIMILARITY.	
DR	-----	DR DISULFD; 805	BY SIMILARITY.
EMBL: AF135253; AAD34456.1; -;	DR DISULFD; 812	BY SIMILARITY.	

		RESULT :5
FT	DISULFID	933 845 BY SIMILARITY.
FT	DISULFID	899 912 BY SIMILARITY.
FT	DISULFID	936 921 BY SIMILARITY.
FT	DISULFID	923 936 BY SIMILARITY.
FT	DISULFID	942 954 BY SIMILARITY.
FT	DISULFID	950 963 BY SIMILARITY.
FT	DISULFID	965 978 BY SIMILARITY.
FT	DISULFID	984 993 BY SIMILARITY.
FT	DISULFID	989 1002 BY SIMILARITY.
FT	DISULFID	1004 1017 BY SIMILARITY.
FT	DISULFID	1023 1035 BY SIMILARITY.
FT	DISULFID	1031 1044 BY SIMILARITY.
FT	DISULFID	1046 1060 BY SIMILARITY.
FT	DISULFID	1066 1079 BY SIMILARITY.
FT	DISULFID	1073 1088 BY SIMILARITY.
FT	DISULFID	1093 1105 BY SIMILARITY.
FT	CARBOHYD	179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	709 755 Missing (in isoform 2).
FT	CONFFLICT	140 159 HSGRKYAAGTHYLSSCRAC -> TVAVSICWPYRPLLP?
FT	CONFFLICT	348 348 GF (IN REF. 2).
FT	CONFFLICT	507 507 S -> L (IN REF. 2).
FT	CONFFLICT	1102 1102 Q -> E (IN REF. 2).
SQ	SEQUENCE	1221 AA; 131818 MW; 87DBB2A10A8FDC45F CRC64;
Qy	Best Local Similarity 10.3%; Score 577.5; DB 1; Length 1221; Matches 157; Conservative 66; Mismatches 182; Indels 137; Gaps 29;	
Db	44 EDVDEAQGLDDOCHADALCQCNPNTSYKC----SCKPGY-QGEGRCEDIDECGNELNGG 97	
Db	768 EDODECLANGTHDGSWKOCFVNLTGSFICVNHTVLCAEGYIINAHRCVDINECVYTLH--766	
Qy	98 C-VHDCLNTPGMYRC---TCDFGMFLAHDHNC-DYDDECLEENNGCC--HTCUNVMGS 145	
Db	767 CTAEHCVNTPTGSFQCQYKALTCEPGYVLT-DG-ECTDVQEVGTGHINCOAGFSQNTKGS 824	
Qy	150 YEC----CCKEGEFLSDNQHTCHRSE-----EGLSCNMKDGCSHLKEAFRGSV 196	
Db	825 FVQCARQRCMDG-FLDDEGNCUDINECTSLLPECRSGFSCN'TVG--SYTCQRND----877	
Qy	197 ACECRGPFLAKHQDCC--IITCNHNGCCOFS--CDDTADGPECSCHPCYKHTDGRSC 252	
Db	878 -LVCGRGHANNEEGSECVDNEETGTHRCGQCLNLPGSYRCOCKPGQRDAGFRGTC 936	
Qy	253 LEREDTVEFTESNTSVYDGDKRYKERLIMETCAVNINGG-CORTCKOTSTGVHSCCPVG 312	
Db	937 ID-----VNECWSGPRLCQHTENTPSYRSCCAAG 968	
Qy	312 FTOOLDGKTKD-DECOTPRNGGDHFCKNIVSSFDGCKKGFKFLTDKSQQDVECSLC 371	
Qy	969 FLLAADDGKHCEDYNECETRR--CSQECAANIXSYQCYCQGQYLAEDGHTCTDECAQ 1026	
Qy	372 R--TCDHSSTINHPGTFAAC--NRGYTLY-GFTHGDIENECSINNGCCQ--QVCVNT-VGSY 425	
Db	1027 AGILCTFRCVNVPDPSYQACPEGGTYMANGRSCKJLDECAALGTHNSSEATCHN-ZGSF 2096	
Qy	426 EC---GCHPGY---KLHNWKDQVEV----KGLL-----KGLL-----449	
Db	1087 RCLRFDCPNEYVRSQTCKERTCQDTCCTSPARITHYCLNFQGLLVPAHFRIGPA 1146	
Qy	450 PTSVSPRVLHCGSKGGDGFCURCHSGIHLSSDVTTTTSYTFKLN--EGKCSLKNAEL 507	
Db	1147 PAFAQDTISLTIK-GNEGYF-----VTRRLNAYTGTVNLSQRSVL 1186	
Qy	508 FP 509	
Db	1167 EP 1.88	

Qy 198 CECRPGFELAKNQRDCILTCNHNGGCQHS CDDTADGPECSCHPOYK2HTGRSCJERED 257
 Db 382 CECKIGYYFDOSRMKV-----DNECQRYP-----GRIC----- 411
 Qy 258 TYLETESNTSWWDGDKRVRRLIMETCAVNNGGDRSKCOTSTGVHSCPVGFTLQLD 317
 Db 412 -----GHK----- 435
 Qy 318 GKTCKDIDECCOTRNGGCDHFCKNIVGSFDCGCKKGFKLJ-TDEKSCCDVDECSL-----DRC 373
 Db 436 GRSCEDINECS -SSPCSQEBANVGSYQCCRGYQLSDGVTCEDIDECAALPTGSHI 493
 Qy 374 CDHSCINHPGTACAC-NRGYTL-YGFTHCCTDNEC-SIMNGGQQVQVNTVGSYEC-- 427
 2b 494 CSYRCINIPGSFOCSGPSSGYRLAPNGRNQDIDECVTGTHKCSINETCFNIQGFRCLIA 553
 Qy 428 -QCHPGYK---LHNWKDKDVE-VKGLLPVS-----PRVSLHCGK--SGGG 467
 Db 554 FCPENYRRSATLQOBKTDVRCIKSCRPNDVTCYDFPVTISHVISLPTFEEFTRSE 613
 Qy 468 DGCFLRCHSGTHLSSDVTIITSVTPKLNES-KCSSLKNAELFPEGL-----RPAJ-P 517
 Db 614 E1FLRAITPPHPAS-----OANIIFDITEGNLRDSFDI-KRYMDGMTVGVVRQVRPIVG 668
 Qy 518 EKHSSYKESFRYV 530
 Db 669 PPHAVLKLEMmrv 681

Search completed: October 22, 2003. 11:28:08
 Job time : 35 secs

PI	Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;	Qy	S41 PGAPGRPPTPKEMITVFELETKETNKEVTTASCDLSCIVKRTLRKAVHREQ 600
PI	Wehrman T, Wang J, Wang D, Drmanac RT;	Db	S41 PGAPGRPPTPKEMITVFELETKETNKEVTTASCDLSCIVKRTLRKAVHREQ 600
XX			
DR	WPI: 2002-759812/82.	Qy	601 PHQJSGMNLDAKKPRTSERQAESCGQGHAENQCSRAGTYDGAERCTLCPNG 663
XX	N-PSDB; ABZ11546.	Db	601 PHQJSGMNLDAKKPRTSERQAESCGQGHAENQCSRAGTYDGAERCTLCPNG 663
PT	New polynucleotides comprising sequences assembled from expressed sequence tags (ESTs) useful for treating cell-proliferative, neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet or coagulation disorders -	Qy	661 TPNNEGGMTCECPRTGSALJKTPBANNSECGLCOPSEADGAPQJCALGTQ 720
PT		Db	661 TPNNEGGMTCECPRTGSALJKTPBANNSECGLCOPSEADGAPQJCALGTQ 720
XX			
PS	SEQ ID NO 1376: 1012pp + Sequence listing: English.	Qy	722 PEAGRITSCFCPCGGLATRKHGATSFQCECTRVQCGSHFNTTHRCCRPGVITYQPEFG 780
XX	The invention relates to an isolated polynucleotide (") comprising a nucleotide sequence selected from any of 948 sequences (ABZ11-9-AB2166) or their mature protein coding portion, active domain coding protein or complementary sequences. The polynucleotides are useful for identifying expressed genes for physical mapping of human genome. The encoded polypeptides (ABP6902-ABP6849) are useful as molecular weight markers, as a food supplement, for generating antibodies, in medical imaging, screening and diagnostic assays and for treating cell-proliferative disorders (cancer), neurodegenerative diseases (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, platelet or coagulation disorders, wound, burns, incisional, ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal, parasitic, arthritis, etc.	Db	721 PEAGRITSCFCPCGGLATRKHGATSFQCECTRVQCSGHFNTTHRCCRPGVITYQPEFG 780
CC	The sequence relates to an isolated polynucleotide ("") comprising a nucleotide sequence selected from any of 948 sequences (ABZ11-9-AB2166) or their mature protein coding portion, active domain coding protein or complementary sequences. The polynucleotides are useful for identifying expressed genes for physical mapping of human genome. The encoded polypeptides (ABP6902-ABP6849) are useful as molecular weight markers, as a food supplement, for generating antibodies, in medical imaging, screening and diagnostic assays and for treating cell-proliferative disorders (cancer), neurodegenerative diseases (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, platelet or coagulation disorders, wound, burns, incisional, ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal, parasitic, arthritis, etc.	Qy	781 KNCVCSGNNTTDFGSTNTITQCKNRGCGELGQFTGIESPQGNYPANTECTWNT 84C
CC		Db	781 KNCVCSGNNTTDFGSTNTITQCKNRGCGELGQFTGIESPQGNYPANTECTWNT 84C
CC		Qy	921 TQPSNEBNNSAARGQVQYVTDQYQELIBDIPVRSRJYASEHNQSLKOKJLKAFLDV 960
CC		Db	921 TQPSNEBNNSAARGQVQYVTDQYQELIBDIPVRSRJYASEHNQSLKOKJLKAFLDV 960
CC		Qy	961 LAHPQNYKTYACESREMPPFSRFLSKYRSFLPYK 999
CC		Db	961 LAHPQNYKTYACESREMPPFSRFLSKYRSFLPYK 999
XX		Qy	961 LAHPQNYKTYACESREMPPFSRFLSKYRSFLPYK 999
SQ	Sequence 999 AA:		
		RESULT 2	
		AB-05567	
		ID ABC05567 standard; protein: 999 AA.	
		XX	
		AC ABZ05567;	
		XX	
		DT 14-NOV-2002 (first entry);	
		XX	
		DE Breast cancer-associated protein 32.	
		XX	
		KW Breast cancer; breast cancer-associated gene sequence; drug development; pharmacogenetics; biosensor development.	
		XX	
		OS Unidentified.	
		PW W0200259377-A2.	
		XX	
		PD 01-AUG-2002.	
		XX	
		PF 24-JAN-2002; 20C2WO-US02242.	
		XX	
		PR 24-JAN-2001; 2001US-263961P.	
		PR 02-FEB-2001; 2001US-265928P.	
		PR 09-APR-2001; 2001US-282698P.	
		PR 04-MAY-2001; 2001US-0829412.	
		PR 29-MAY-2001; 2001US-288590P.	
		XX	
		(ECSB-) EOS BIOTECHNOLOGY INC.	
		XX	
		PW Yack CH, Gish KC, Afar D;	
		XX	
		WP1: 2002-562738-62.	
		DR N-PSDB; ABZ07724.	
		XX	
		PR Detecting a breast cancer-associated transcript in a patient's cell, comprising contacting a biologic sample with a polynucleotide that selectively hybridizes with breast cancer nucleic acids -	
		XX	

PS Disclosure: Page 375; 414pp; English.

XX The invention comprises a method of detecting a breast cancer-associated transcript in a patient. The method of the invention involves contacting a biological sample from the patient with a nucleotide that hybridises to one of the 69 breast cancer-associated gene sequences shown in the specification. The method of the invention is useful in the diagnosis or prognosis of breast cancer, and for detecting genes that are up or down regulated in breast cancer cells. Genes identified by the method of the invention can be used in diagnostic purposes and also as targets for screening for therapeutic compounds that modulate breast cancer (e.g. hormones or antibiotics). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. Amino acid sequences ABJ05536 - ABJ0564 represent the proteins encoded by the 69 breast cancer-associated genes of the invention.

XX Sequence 999 AA;

Query Match 100.0%; Score 5605; DB 23; Length 999;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGYAGRRPGAAWAVLILLLPPLILLAGVPPGRRAACPAQDQEYDECAQGLDDCHADA 6C
Db 1 MGYAGRRPGAAWAVLILLLPPLILLAGVPPGRRAACPAQDQEYDECAQGLDDCHADA 6C
Qy 61 LCONTNPSSYKCSCKPGYQGBQRCEDIDEQNENLNGCVHDLNIPGNYTROTCPDFGMLA 120
Db 61 LCONTNPSSYKCSCKPGYQGBQRCEDIDEQNENLNGCVHDLNIPGNYTROTCPDFGMLA 120
Qy 121 HDGHNCNLDVDCELENNGCQHTCVAHSEEGHSCKMK 180
Db 121 HDGHNCNLDVDCELENNGCQHTCVAHSEEGHSCKMK 180
Qy 181 DHGCSHICKEAPRGSYACECRGFELAKNORDCILTONHNGNGCQHSCDDTDAGPECSCH 240
Db 181 DHGCSHICKEAPRGSYACECRGFELAKNORDCILTONHNGNGCQHSCDDTDAGPECSCH 240
Qy 241 POKYKMTDGRSCLEREDTVLETESENNTSDGDKEVKRKRLMETCAVNNGCDRTCKT 30C
Db 241 POKYKMTDGRSCLEREDTVLETESENNTSDGDKEVKRKRLMETCAVNNGCDRTCKT 300
Qy 301 STGVHCSOPGVFTQLDQGKTCIDECQTRNGCDHFCKNTIVGSFDQ3CKGFKLTDK 360
Db 301 STGVHCSOPGVFTQLDQGKTCIDECQTRNGCDHFCKNTIVGSFDQ3CKGFKLTDK 360
Qy 36: SCQDVDESLDRTCDHSCKINHPGTFACACNRGCTLYGFTHGDTNECSNNGGCCQCYCN 420
Db 361 SCQDVDESLDRTCDHSCKINHPGTFACACNRGCTLYGFTHGDTNECSNNGGCCQCYCN 420
Qy 421 TVGSYECPHPGYKLHNKKOVCVEKGKLGPTSPVLSIHQGKSGGDSCFLRCHSGIHL 480
Db 421 TVGSYECPHPGYKLHNKKOVCVEKGKLGPTSPVLSIHQGKSGGDSCFLRCHSGIHL 480
Qy 481 SDDVTITRTSTVKLNKGCKSLNAELFPEGRLPERKSSKESFRYVNLTCSSGKQV 540
Db 481 SDDVTITRTSTVKLNKGCKSLNAELFPEGRLPERKSSKESFRYVNLTCSSGKQV 540
Qy 54: PGAPGRPSTPKMFMITYFEELTNOKEVTASCDLSLCKVTEKLRKIAITRKAVHREQ 60C
Db 541 PGAPGRPSTPKMFMITYFEELTNOKEVTASCDLSLCKVTEKLRKIAITRKAVHREQ 60C
Qy 601 FHQLQSGNLDVAKKPRPTSEROAESCGVGQGHAENQCVSRAGTYDGARERCILCPNS 66C
Db 601 FHQLQSGNLDVAKKPRPTSEROAESCGVGQGHAENQCVSRAGTYDGARERCILCPNS 66C
Qy 661 TFQNEEGMTCTBPCPRPONSGALKTPAWNSECGLICQPBEYSDGFPQLQALGTQ 720
Db 661 TFQNEEGMTCTBPCPRPONSGALKTPAWNSECGLICQPBEYSDGFPQLQALGTQ 720

Qy 721 PEAGRITSSCFPCGGGATKHQAGATSPQDCETRVQCSQPHFYNNTTHRCIRCPVGTVOPEFG 780
Db 72: PEAGRITSSCFPCGGGATKHQAGATSPQDCETRVQCSQPHFYNNTTHRCIRCPVGTVOPEFG 78C
Qy 78: KNCVUSCGNNTTDFGDSNTIQCQKRCRGGEJGDFTSYIESPNYPOVNPANTECWTIN 84C
Db 781 KNCVUSCGNNTTDFGDSNTIQCQKRCRGGEJGDFTSYIESPNYPOVNPANTECWTIN 84C
Qy 841 PPKRTRIVVPPIFPEEDGQYQYMRKTSSSNSVTTYETCQYERPIATTSRSKLV 900
Db 841 PPKRTRIVVPPIFPEEDGQYQYMRKTSSSNSVTTYETCQYERPIATTSRSKLV 900C
Qy 901 TQFKSNENSARGSQFCVYTYDDEYQELIEDIVDRGSRLYASENHGEELDKKLIXAIFCV 96C
Db 901 TQFKSNENSARFCVYTYDDEYQELIEDIVDRGSRLYASENHGEELDKKLIXAIFCV 96C
Qy 961 LAHPONYKYTKYTAQESREYFPRSFIRLRSKVSRELRYK 99
Db 961 LAHPONYKYTKYTAQESREMFPRSFIRLRSKVSRELRYK 999

RESULT 3
ABB80926 standard; Protein: 999 AA.
XX
AC ABB80926;
XX DT 08-OCT-2002 (first entry)
XX DE Human breast cancer modulating protein BCC2.
XX KW Breast cancer modulating protein; BCMP; BCO2; cyostatic; human;
KW immunosimulant; antisense therapy; gene therapy; vaccine.
XX Homo sapiens.
XX PN KO2C255988-A2.
XX PR 18-JU-2002.
XX PR 21-DEC-2001; 2001WO-US50548.
XX PR 21-DEC-2000; 2000US-C747371.
XX PR (E2SB-) ; ECS BIOTECHNOLOGY INC.
XX Gish KC, Mack DJ;
XX NP-; 2002-566249/60.
XX N-2SDB; ABN96363.
XX PT Diagnosing breast cancer comprises comparing the gene expression of a breast cancer modulating protein in a test breast tissue and unaffected breast tissue of the same or a different patient -
XX Disclosure: Fig 2; 68pp; English.
XX
The invention relates to diagnosing breast cancer that involves determining the expression of a gene encoding breast cancer modulating protein (BCO2) or a fragment of it in breast tissue of an individual, and comparing the expression of the genes from a second normal breast tissue from the individual or a second unaffected individual, where a difference in expression indicates the individual has breast cancer. The method is used for diagnosing breast cancer. Other new methods are provided for screening drug candidates, screening for bioactive agents, and evaluating the effect of a candidate breast cancer drug. An antibody to BCO2 can be used to inhibit the bioactivity of BCO2 or neutralise its effect and can be used to treat breast cancer. The antibody can also be used to localise a therapeutic group to breast cancer tissue, which can be used to treat breast cancer. Antisense molecules are used to inhibit breast cancer in a cell. BCO2 protein and polynucleotides are used to elicit an immune response in an individual. BCO2 is also used to determine the prognosis of an individual with breast cancer. The

Db	1 NGVAGNRPGAAWAEVLLILPPPLLAGVPPGRRAAQPQDVDECAQGLDCHADA	60	XX	12-JUN-2003 (first entry)
Qy	61 LCCNTPSYKSCCKPQGYGEGRCEDIDECCNELNQCVHDCLNIPGNYRCTCFDGFMLA	123	DE	Breast cancer associated protein sequence SEQ ID NO:45.
Db	61 LCCNTPSYKSCCKPQGYGEGRCEDIDECCNELNQCVHDCLNIPGNYRCTCFDGFMLA	120	XX	
Qy	121 HDGHNCNLDVDCELENNGCQHTCIVNNGSYECCCKESEFFSLDSNQHTC1HSEEGLSCKNK	180	XX	Human; breast cancer; cytosatic; gene therapy.
Db	121 HDGHNCNLDVDCELENNGCQHTCIVNNGSYECCCKESEFFSLDSNQHTC1HSEEGLSCKNK	180	OS	Homo sapiens.
Qy	181 DHGCSHCKEAPRSVACECPGFELAKQDC1LTC1HNGGCCDHSCTADGPECSCH	240	PK	WO200304999-A2.
Db	181 DHGCSHCKEAPRSVACECPGFELAKQDC1LTC1HNGGCCDHSCTADGPECSCH	240	XX	16-JAN-2003.
Db	21 DHGCSHCKEAPRSVACECPGFELAKQDC1LTC1HNGGCCDHSCTADGPECSCH	240	PF	21-JUN-2002; 2002WO-US19663.
Qy	241 POKMHDGRSCLEREDTVLETESNTTISVYGDKRVKRRLMETCAVNINGGCDPRTCKDT	300	XX	21-JUN-2001; 2001US-299887P.
Db	241 POKMHDGRSCLEREDTVLETESNTTISVYGDKRVKRRLMETCAVNINGGCDPRTCKDT	300	PR	27-JUN-2001; 2001US-301572P.
Qy	301 STGVHSCSPCPVGETLQDGKTKDIDCQTRNGCQDFHFCRN1VGSFDCCGKKGKFLLTDK	362	PR	18-JUL-2001; 2001US-306501P.
Db	301 STGVHSCSPCPVGETLQDGKTKDIDCQTRNGCQDFHFCRN1VGSFDCCGKKGKFLLTDK	360	PR	25-SEP-2001; 2001US-3250C2P.
Qy	361 SCQDVDECSDLPCTDHSICINHPGTFACAGNGTYLYGFTRGDNECSINNGGQQQCVYN	420	PR	05-MAR-2002; 2002US-362589P.
Db	361 SCQDVDECSDLPCTDHSICINHPGTFACAGNGTYLYGFTRGDNECSINNGGQQQCVYN	420	XX	14-MAY-2002; 2002US-380391P.
Qy	421 TGSYEQCHPSKYLHNWKDQCVKGLLPTSVSPVSLHGKSGGDGTFLRCHSGIHL	480	PA	(MIL-1) MILLENIC PHARM INC.
Db	421 TGSYEQCHPSKYLHNWKDQCVKGLLPTSVSPVSLHGKSGGDGTFLRCHSGIHL	480	XX	Lillie J., Gannavarapu M., Glatt K., Hoersch S., Kamatkar S., Mertens N.; Monahan JE., Myer V., Wang Y., Xu Y., Zhai X., Meyers RE., East RC.; Hertogoyi GN., Puszta L., Sahn A., Milić GB;
Qy	481 SSDVTLRTSTFKLNEGKCSLKAELPEGLRPAKHSVSKESFRYVNLTCSSGKV	540	XX	WPI: 2003-2-0381/20.
Db	481 SSDVTLRTSTFKLNEGKCSLKAELPEGLRPAKHSVSKESFRYVNLTCSSGKV	540	DR	N-PSDB; ACCS0098.
Qy	541 PGAPGRPSTPKEMFITFEELTNOKEYTASCDLSCIVKRTKLRKAI1T1RKAVHREQ	600	XX	Breast cancer diagnosis or treatment by comparing the level of expression of a marker in a patient sample with that in the control non-breast cancer sample -
Db	541 PGAPGRPSTPKEMFITFEELTNOKEYTASCDLSCIVKRTKLRKAI1T1RKAVHREQ	600	PS	Claim 1; SEQ ID 45; 128pp; English.
Qy	601 FHIQLSGMMNLDVAKKPRTSBRQAESCGVGQGHAENQCVSFRAGTYDGAERCHJCPNG	660	CC	The present invention describes a method for assessing whether a patient is afflicted with breast cancer. The method comprises comparing the level of expression of a marker [gene/po-peptide see ACCS0098 to ACCS034 and AB47386 to ABR4632] in a patient sample and the normal level of expression of the marker in a control non-breast cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with breast cancer. The breast cancer associated sequences from the present invention have cytostatic activities and can be used in gene therapy. The method is useful for diagnosing and treating breast cancer.
Db	601 FHIQLSGMMNLDVAKKPRTSBRQAESCGVGQGHAENQCVSFRAGTYDGAERCHJCPNG	660	CC	K.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pdb/published_pct_sequences .
Qy	661 TFOONEEGOMTCBPCPRPGNSGALKTPEAMNSCECGGLCOPGEYSADGFAPOQCLALGSTFO	720	XX	Sequence 999 AA:
Db	661 TFOONEEGOMTCBPCPRNGSALKTPEAMNSCECGGLCOPGEYSADGFAPOQCLALGSTFO	720	CC	Query Match 100.3%; Score 5605; DB 24; Length: 399;
Qy	721 PEARGTSCKPCGGGLATKHOAGTSFQDCETVRQCSGPQHFMNTTHRC1RPVGTQYQPEFG	780	Qy	Best Local Similarity: 100.3%; Pred. No.: 0; Mismatches: 0; Indels: 0; Gaps: 0;
Db	721 PEARGTSCKPCGGGLATKHOAGTSFQDCETVRQCSGPQHFMNTTHRC1RPVGTQYQPEFG	780	Db	Matches: 399; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
Qy	781 KDNCVSPGNNTTDFDOSTNTITQCKNRRCGELGDFGTG1IESPNPGYVANTECTWTIN	840	Qy	2 MGVAGRNPQGAAWAVLILPPPLIAGAAPPGRGRAGQEDVDECAQGDDCHADA 60
Db	781 KDNCVSPGNNTTDFDOSTNTITQCKNRRCGELGDFGTG1IESPNPGYVANTECTWTIN	840	Db	1 MGVAGRNPQGAAWAVLILPPPLIAGAAPPGRGRAGQEDVDECAQGDDCHADA 60
Qy	841 PPKRRLILIVBIFLIEDDGDYLMRKISSNSVTTYTCQYERPAFTSRSSKKLW	900	Qy	61 LCONCPTSYKSCSPKGEGSQCFDIECCNELNQCVHDCLNIPGNYRCTCFDGFMLA 120
Db	841 PPKRRLILIVBIFLIEDDGDYLMRKISSNSVTTYTCQYERPAFTSRSSKKLW	900	Db	61 LCONCPTSYKSCSPKGEGSQCFDIECCNELNQCVHDCLNIPGNYRCTCFDGFMLA 120
Qy	901 IQFKSNEGNSARGFQPVYTYDQYQELIEDIVDRGRLYASENHOI1DKKLKALFV	960	Qy	121 HOGHNCLDVECLENKGCOHQTCVNNNGSYECCCKESEFFLSDNQHTJHRSEBGLCMNK 180
Db	901 IQFKSNEGNSARGFQPVYTYDQYQELIEDIVDRGRLYASENHOI1DKKLKALFV	960	Db	121 HOGHNCLDVECLENKGCOHQTCVNNNGSYECCCKESEFFLSDNQHTJHRSEBGLCMNK 180
Qy	961 LAHPONYKTYTAQESREMFRSFIRLRSKYSRFLPYK 999	999	Qy	181 DHGCSHCKEAPRSVACECPGFELAKNORDCILTC1HNGGCCSDDTAQGPSCCH 240
Db	961 LAHPONYKTYTAQESREMFRSFIRLRSKYSRFLPYK 999	999	Db	181 DHGCSHCKEAPRSVACECPGFELAKNORDCILTC1HNGGCCSDDTAQGPSCCH 240
RESULT				
ABR47407				
ID				
XX				
AC				

Qy	241	PQYKMHHTDGPSCLEPFDTYLAVTEESNTTSVUDGKVRKPLIMETCAVNGGCDRTCKT	305	PF 04-JUN-2002; 2002WO-US17594.
Db	241	PQYKMHHTDGPSCLEPFDTYLAVTEESNTTSVUDGKVRKPLIMETCAVNGGCDRTCKT	305	PR 04-JUN-2001; 2001US-2959-7P.
Db	291	STGVHCSCPVGYFTLQDGKTKDIEBCQTRNGGCDHFCKNIVGSFDCGKKGFKLTLDEK	362	PR 13-NOV-2001; 2001US-350668P.
Qy	301	STGVHCSCPVGYFTLQDGKTKDIEBCQTRNGGCDHFCKNIVGSFDCGKKGFKLTLDEK	362	PR 29-MAR-2002; 2002US-168689P.
Db	301	STGVHCSCPVGYFTLQDGKTKDIEBCQTRNGGCDHFCKNIVGSFDCGKKGFKLTLDEK	362	PR 12-APR-2002; 2002US-37246P.
Db	301	STGVHCSCPVGYFTLQDGKTKDIEBCQTRNGGCDHFCKNIVGSFDCGKKGFKLTLDEK	362	PR 31-MAY-2002; 2002US-016C233.
PA		(EOSB-) EDS BIOTECHNOLOGY INC.		XX PA
Qy	361	SCQDYBECSSLRDTCDHSINHPGTFAACANRGTYLYFGTHCDPNECSINNGGQQVCVN	420	XX PR Detecting an androgen-independent prostate cancer cell in a sample or
Db	361	SCQDYBECSSLRDTCDHSINHPGTFAACANRGTYLYFGTHCDPNECSINNGGQQVCVN	420	XX PR diagnosing androgen-dependent prostate cancer, by determining the presence or absence of genes whose expressions are up- or down-regulated.
Qy	421	TVGSYECCHGPGYKLHWNKDKCVDVEYKGLIPLTSVSPRUSIHCKGSSGGCFCFLRCHSGTHL	480	XX PR
Db	421	TVGSYECCHGPGYKLHWNKDKCVDVEYKGLIPLTSVSPRUSIHCKGSSGGCFCFLRCHSGTHL	480	XX PR
Qy	481	SSDCYTIRTSYTFKLNEGKCSLKAALFPECLRPLPEKHSSKESFRVNLTCSSGKV	540	XX PR
Qy	481	SSDCYTIRTSYTFKLNEGKCSLKAALFPECLRPLPEKHSSKESFRVNLTCSSGKV	540	XX PR
Db	481	SSDCYTIRTSYTFKLNEGKCSLKAALFPECLRPLPEKHSSKESFRVNLTCSSGKV	540	XX PR
Qy	541	EGAPGRPSTPKEMFTIVEFEDELTNKEVTAASCDLSIVKTKERJKAIRTLKAVHREQ	600	XX PS Claim 1; Page 199; 2103P; English.
Db	541	EGAPGRPSTPKEMFTIVEFEDELTNKEVTAASCDLSIVKTKERJKAIRTLKAVHREQ	600	XX PS The invention comprises a method for detecting an androgen-independent cancer cell in a sample from a patient who has undergone androgen ablation therapy. The method involves determining the presence or absence of nucleic acids that are either up-regulated or down-regulated in prostate cancer. The method is useful for detecting an androgen-independent prostate cancer cell in a sample from a patient who has undergone androgen ablation therapy. The method is particularly useful for diagnosing androgen-dependent prostate cancer, prostate cancer undergoing androgen withdrawal, or androgen-independent prostate cancer. A gene that is either up-regulated or down-regulated in prostate cancer.
Qy	601	FHLQLGSMNLVAKKPPRTSERGAFSCGVQSGHASNQCSRACTYYDGARFCLCPNG	665	XX PS
Db	601	FHLQLGSMNLVAKKPPRTSERGAFSCGVQSGHASNQCSRACTYYDGARFCLCPNG	665	XX PS
Qy	661	TFCQNEQMTCEPCCPQGNSSALKTPBAMNNSECGGLCQPEYEYSDGAPCQLGALGFPQ	720	XX PS
Db	661	TFCQNEQMTCEPCCPQGNSSALKTPBAMNNSECGGLCQPEYEYSDGAPCQLGALGFPQ	720	XX PS
Qy	721	PEAGRITSCFPCGGGLATKHOATSFDQCETRVQCPGHEYNNTTHRCZRPVG-YQPERG	780	XX PS
Db	721	PEAGRITSCFPCGGGLATKHOATSFDQCETRVQCPGHEYNNTTHRCZRPVG-YQPERG	780	XX PS
Qy	781	KNNCVSPGNNTTDFGSTNTTQCKNVRRCGELGDFDTGYIESPNYPGNYPANTECWLN	840	XX PS
Db	781	KNNCVSPGNNTTDFGSTNTTQCKNVRRCGELGDFDTGYIESPNYPGNYPANTECWLN	840	XX PS
Qy	841	PPPKRRLIVVPEIFLPIEDDGDYLVMRKSSNSVTTYETCQYERPIAFTSRSKLW	900	XX PS
Db	841	PPPKRRLIVVPEIFLPIEDDGDYLVMRKSSNSVTTYETCQYERPIAFTSRSKLW	900	XX PS
Qy	901	IQFKSNEGNNSARGFQPYVTYDEDYQBLIEDIVDRGLYASENHQEILDKKLKALFDY	960	XX PS
Db	901	IQFKSNEGNNSARGFQPYVTYDEDYQBLIEDIVDRGLYASENHQEILDKKLKALFDY	960	XX PS
Qy	961	LAHPNQNYFKYTAQESREMPPSFIRLRSKYSRFLRPyK	999	XX PS
Db	961	LAHPNQNYFKYTAQESREMPPSFIRLRSKYSRFLRPyK	999	XX PS
RESUL	6			
ID	ABJ19815	standard; Protein: 999 AA.		
AC	ABJ19815;			
DT	10-APR-2003	(first entry)		
XX		Androgen-independent prostate cancer-related protein - SEQ ID No 42.		
KW		Androgen-independent cancer; androgen ablation therapy; prostate cancer.		
KW		androgen-dependent prostate cancer; prostate cancer.		
XX		Unidentified.		
CS				
DE				
XX				
AC				
XX				
DT				
XX				
XX				
PN	WO200298358-A2.			
XX				
PD	12-DEC-20C2.			
XX				

Qy	961	LAHPOQNYFKYTAQESREMPSFIRLRSKVSRELRYK	999	
Db	959	LAHPQNYFKYTAQESREMPSFIRLRSKVSRELRYK	997	
		RESULT 8		
	ABP56757	ID ABP56757 standard; Protein: 964 AA.		
XX				
AC	ABP56757;			
XX				
DT	31-MAR-2003	(first entry)		
XX				
DE	Human CEGP1 protein SEQ ID NO:5.			
XX				
KW	secreted protein; epidermal growth factor protein subfamily;			
KW	epidermal growth factor; EGF; gene therapy; therapeutic; drug screening;			
KW	biological; activity; immune response.			
XX				
OS	Homo sapiens.			
XX				
PN	NC20C210080-A2.			
XX				
PD	19-DEC-2002.			
XX				
PF	07-MAY-2002; 2002M0-US222278.			
XX				
PR	16-MAY-2001; 2001US-0855824.			
XX				
PA	{PEKE } PE CORP NY.			
PA	(GONG/) GONG F.			
PA	(DFRA/) DI FRANCESCO V.			
P1	Gong F., Di Francesco V., Yan C., Beasley EM;			
XX				
PT	2003-167407/16.			
XX				
PS	Disclosure; Fig 2E-F; 185pp; English.			
XX				
CC	The present sequence encodes a human secreted peptide (I) related to			
CC	the epidermal growth factor (EGF) protein subfamily. (I) can be used			
CC	in gene therapy. (I) and the nucleic acid molecules encoding (I) can			
CC	be used as models for the development of human therapeutic targets, to			
CC	aid in the identification of therapeutic proteins, and to serve as			
CC	targets for the development of human therapeutic agents. (I) may be			
CC	used in drug screening assays, in assays to determine the biological			
CC	activity of the protein, or to raise antibodies or to elicit another immune			
CC	response, as a reagent in assays designed to quantitatively determine			
CC	levels of the protein in biological fluids, or as markers for tissues			
CC	in which the corresponding protein is preferentially expressed. (I) can			
CC	also be used for diagnosing or treating a disease or disorder			
CC	characterised by an absence of, inappropriate or unwanted expression of			
CC	the protein. (I) is located to human chromosome 22. The present sequence			
CC	represents human CEGP1 protein which is given in comparison with (I)			
XX	in the exemplification of the present invention.			
SQ	Sequence 964 AA.			
Qy	Query Match Score 96.8%; Score 5428; DB 24; Length 964;			
	Best Local Similarity 100.0%; Pred. No. 0;			
	Mismatches 0; Indels 0; Gaps 0;			
	Yatches 964; Conservative 0;			
Db	36 RGRAGPQEDVDECAQGLDDCHADALCONTPTSYKCSCKPGYGEQCEJDICONELN 95			
,	1 RGRAGPQEDVDECAQGLDDCHADALCONTPTSYKCSCKPGYGEQCEJDICONELN 60			
Qy	96 GGCVHDCLNIPGNYRCTGDFGMЛАДГНСЛДВЕДЛЕНГСОИТСВИМГСҮЕССК 155			

RESULT 3

AAC16642

AAC16642 standard; Protein: 919 AA.

XX

AC

XX

AC

XX

AC16642;

XX

AC

KW Human; gene therapy; protein replacement therapy;
 extracellular messenger; EXMES; cardiovascular disorder; cancer;
 KW congestive heart failure; autoimmune disorder; inflammatory disorder;
 AIDS; neurological disorder; epilepsy; reproductive disorder;
 infertility; infection; developmental disorder; Cushing's syndrome;
 endocrine disorder; diabetes insipidus; cell proliferative disorder.
 XX
 Homo sapiens.
 XX
 PN WO2003002610-A1.
 XX
 09-JAN-2003.
 PD
 XX
 PF
 XX
 PR 29-JUN-2001; 2001US-301789P.
 PR 21-SEP-2001; 2001US-324149P.
 PR 05-OCT-2001; 2001US-327713P.
 PR 12-OCT-2001; 2001US-329215P.
 PR 14-DEC-2001; 2001US-340218P.
 PR 05-APR-2002; 2002US-370761P.
 PR 19-APR-2002; 2002US-373824P.
 XX
 (INCYT) INCYTE GENOMICS INC.
 PA Deleagean AM, Borowsky ML, Khan FA, Kearney L, Ramkumar J;
 PI Khan NR, Lu Y, Honchell CD, Emerling BM, Gorvad AE;
 PI Griffin JA, Warren BA, Yue H, Thangavelu K, Sprague WW, Ison CH;
 PI Elliott VS, Mason PM, Richardson TN, Tran UK, Swarnakar A, Jin P;
 PI Kable AE;
 XX
 DR 2003-210247/20.
 DR N-PSDB; AALS1944.
 XX
 PT New human extracellular messengers (EXMES) and polynucleotides, useful
 for diagnosing, treating or preventing e.g. congestive heart failure,
 AIDS, stroke, Alzheimer's disease, infertility, infections, leukemia or
 breast cancer -
 XX
 PS Claim 1; Page 182-184; 243PP; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 extracellular messengers (EXMES). The EXMES DNA and protein sequences are
 useful for treating a disease or condition associated with decreased/
 increased expression of functional EXMES. The EXMES DNA and protein
 sequences are useful for diagnosing, treating and preventing:
 CC cardiovascular disorders (e.g. congestive heart failure); autoimmune/
 inflammatory disorders (e.g. AIDS); neurological disorders (e.g.
 epilepsy); reproductive disorders (e.g. infertility); infections;
 development disorders (e.g. Cushing's syndrome); endocrine disorders
 (e.g. diabetes insipidus); and cell proliferative disorders (e.g.
 cancer). The present amino acid sequence represents a human EXMES protein.
 CC
 Sequence 919 AA;

Query Match 90.6%; Score 5079; DB 24; length 919;
 Best Local Similarity 91.7%; Pred. No. 2.7e-307;
 Matches 916; Conservative 1; Mismatches 2; Indels 80; Gaps 1;

Qy 1 MGIVAGRNRPGAAVAVLILLIPLLILLIAGAVPPGRPAAGGQEDVDECAQLDCHADA 60

Db 1 MGIVAGRNRPGAAVAVLILLIPLLILLIAGAVPPGRPAAGGQEDVDECAQLDCHADA 60

Qy 6 LCQNTPTSYKCSCKPQYQEGEGROEDIDEDECGNEELNGGVHDC-NIPGNYRCFCDFGMLA 120

Db 61 LCQNTPTSYKCSCKPQYQEGEGROEDIDEDECGNEELNGGVHDC-NIPGNYRCFCDFGMLA 120

Qy 121 HDGHNCLDVECLENNGGQHTCUNMSYECCKEGFFLISDNQHTCHRSEFGSCMNK 180

Db 122 HDGHNCLDVECLENNGGQHTCUNMSYECCKEGFFLISDNQHTCHRSEFGSCMNK 180

Qy 181 DHGSHICKEAPRSVACECRPFELAKNQRCILTCNGNGCQHSDDTAGPESCH 240

Db 181 DHGSHICKEAPRSVACECRPFELAKNQRCILTCNGNGCQHSDDTAGPESCH 240

CY 241 POKXHTDGRSCLEREDTYLETESNTTSVVDGDKVRRLIMETCAVNNGCORTKD 300

Db 242 POKXHTDGRSCLEREDTYLETESNTTSVVDGDKVRRLIMETCAVNNGCORTKD 300

CY 3C1 STGYACSPVGFHQGLQKTKCDIDECCTRNGCDECKCQVGSFDGCKKGPKLTDK 360

Db 3C2 STGYACSPVGFHQGLQKTKCDIDECCTRNGCDECKCQVGSFDGCKKGPKLTDK 360

CY 361 SCQDV2ECSLDRTDCHSCINHPGTACACNRGVTLYGTHCCDNTNECSSINNG3CQCVCN 420

Db 361 SCQDV2ECSLDRTDCHSCINHPGTACACNRGVTLYGTHCCDNTNECSSINNG3CQCVCN 420

CY 421 TVGSYECCQCHPGYKLHWNKKDCVEVKGILP7SVSPRVSLHGCKSGGCCPFRCHSGH 480

Db 403 ----- 432

CY 481 SSDVTTIRTSVTFKLNEGKCSL-KNAELPEGLRPAJPEKHSSVKESEPRYYNL-CSSGKV 540

Db 403 --CVTTIRTSVTFKLNEGKCSL-KNAELPEGLRPAJPEKHSSVKESEPRYYNL-CSSGKV 460

CY 541 PGAGRPSPPKENFITEFELEINQENTASCDLSCLVTRKRAIRTEKAVERBQ 600

Db 461 PGAGRPSPPKENFITEFELEINQENTASCDLSCLVTRKRAIRTEKAVERBQ 520

CY 6C1 FHQLSGMNLQVAKKPPRTSEROAESCCVGGHAENOCVSRAGTYYGARERCILCPNG 660

Db 522 FHQLSGMNLQVAKKPPRTSEROAESCCVGGHAENOCVSRAGTYYGARERCILCPNG 580

CY 661 TFOQNEBEGMTCECP-PRPNSNMGALKTPPEAWNMSPCGGLCOPGEYSACGFAPCQLCALGTFQ 720

Db 581 TFOQNEBEGMTCECP-PRPNSNMGALKTPPEAWNMSPCGGLCOPGEYSACGFAPCQLCALGTFQ 640

CY 721 PEAGRTSCPCCGGSLATKHGATSFQDGTTRVQCSPHFYNTTHRCCRPMGTYQPERG 780

Db 641 PEAGRTSCPCCGGSLATKHGATSFQDGTTRVQCSPHFYNTTHRCCRPMGTYQPERG 700

CY 781 KNKNYCSPEPNTTDEDGSTNITOCKNRGEGLDDFTYIESPNYPGNYPANTECTMIN 840

Db 701 KNKNYCSPEPNTTDEDGSTNITOCKNRGEGLDDFTYIESPNYPGNYPANTECTMIN 760

CY 842 PPPEKRIL-VYPE-FLPLDEDCCSDYLVMRKCSSSNNSVTVETCYEPFAFTSRSKLW 900

Db 761 PPPEKRIL-VYPE-FLPLDEDCCSDYLVMRKCSSSNNSVTVETCYEPFAFTSRSKLW 920

CY 901 IQFKSNEGNSARGQVQPVYTYDQYQELIED-YVDRGRLYASENHOE-LKOKKLKALFDV 960

Db 821 IQFKSNEGNSARGQVQPVYTYDQYQELIED-YVDRGRLYASENHOE-LKOKKLKALFDV 980

CY 961 LAHQNYFKYTAQESRENPERSFIRLGSKVSREPRYK 999

Db 881 LAHQNYFKYTAQESRENPERSFIRLGSKVSREPRYK 919

RESULT

ABBE-327 standard; Protein: 997 AA.

AC ABBB-327;

DC 06-OCT-2002 (first entry)

MC use BC02 orthologue protein.

Breast cancer modulating protein; BC02; cytostatic; mouse;
 immunostimulant; antisense therapy; gene therapy; vaccine;
 Xus sp.

WC200255988-A2.

XK

Best Local Similarity 89.2%; Pred. No. 8e-306; Length 39; Matches 891; Conservative 47; Mismatches 59; Index 2;

1 MGVAAGRNRPGAAWAVLILLPPPLLAGAVPPGRGRAAGFPQEDVDECAQGLY

卷之三

...MUGAUGURKPEAKALLLULPPF - LUMAAVPEURGLINGFSEUDVLEAGL

61 CCNTPTSYKCSCSKPBYOGEGROCEDIRECGNELNGGCYHDCLNIPGNYBCC

卷之三

b 59 LCONTPTSYKCSCKPGYQGEGRQCEDMDECNDTLNGGCVHDCLNIPGMYRCTC

Y Y
121 HBGHNCLBVBECLNNNGCCAHTEVNMGSYECCKEGFFLJ

卷之三

THE INFLUENCE OF CULTURAL INVOLVEMENT ON FUSION

1 81 DHGCSHICKEAPRGSSVACECBRGFEELAKNORDCILTCN4GNGCCOHSSCDTAD

卷之三

b b 279 DHGCGHICKEAPRGSVACECRPGFELAKNOKDCILTCNHNGGGCOHSCECTAE

241 PQYKMHDTGRSCLEREDTVLEVTESENTTSVVDGGDKRVKRRLLMETCAVNNGGC

卷之三

233 FRIKIRHADGRSCLLEQES: VLEGIESNA: SVADGDURRVERLIME: CAVNNGSC:

3.01 SITEVHCCSPYGETLQIDGKTCKDIDECOTBNGGCCHECKNIVVSEDCCCCC

卷之三

299 STGWHCSCTPGFTLQVDGKTKDIDECOTRNGCNHECKNTVGSFDCCSKKGFB

361 SCQDVDECSDLRTCDHSCINHPGTFACACNRGTYLGYCFTHCDDTNESCSINNGG
Y

卷之三

SCQBVDECSRERTCDHSCLINERPGTIFICACNPGYILYSSFHCGDBNECSVNGG

421 *TEASEREFERENCEHACKERIAHUMANOMONOCYCI*: PREVIEWPAGE. HUMANOMONOCYCI

卷之三

Qy	481	SSDVYVITRTSVTFKINNEGKCSLKNALPEFGLRPLPDXHSSVYKESTRYVNLTCSSSKQV	540
Db	479	SSDVYVITRTSVTFKINNEGKCSLQXAKLSPGLPAPLPERHSSVYKESTQYANLTCSPSKQV	538
Qy	541	PGAPGRSPTEKEMITVEPEJETNOKEYTASCDLSCLVKTRERKLRAIRTLRKAVREO	600
Db	539	PGALCRLNAPKEMITVEPEJETVEKEVAAASCNJSCUVKRTBKLRLTRKLRAIREQ	598
Qy	601	FHLC-SGMN-LDVAKKPRSERGAESCGVGQGHAEANGQVSCHAGTYDGRERCLCPNG	660
Db	599	FHLO-SGMOD-LDVAKKPRSERGAESCGVGQGHAEANGQVSCHAGTYDGRERCLCPNG	659
Qy	661	TFCNEEGQNTCEPDPFGNSGALKTPEAMNMSEGGLGQCGEYSAODGAPCQLCALGTFO	723
Db	659	TFCNEEGQNTCEPDPFGNSGALKTPEAMNMSEGGLGQCGEYSAODGAPCQLCALGTFO	718
Qy	721	PEAGEFSCPCGGGLATKHQGATSFQDCETTRVOCSPGFYNTTTCRCRCPVSTYQPBEFG	780
Db	719	PEAGEFSCUSCGGGGLPTKHQGATSFQDCETTRVOCSPGFYNTTTCRCRCPVSTYQPBEFG	778
Qy	781	KNNCYSCGEHTTTDGSNTITQCKNRKGGLGDFGTGYIESPNPQNPYPAANTECTWNT	840
Db	779	KNNCYSCGEHTTTDGSNTITQCKNRKGGLGDFGTGYIESPNPQNPYPAANEBCTWNT	838
Qy	841	PPPKERLIVWPEIPLPIBDGCDYLWMRKTSSNSVTYETCTYBZTAFTSRSKCKW	900
Db	839	PPPKERLIVWPEIPLPIBDGCDYLWMRKTSSNSVTYETCTYBZTAFTSRSKCKW	898
Qy	901	IQFKSNEGSARGCQPVYTYDDEQJELIEDVDRGLYASENHCEILKCKLKLKA.FDV	960
Db	899	IQFKENEGSARGCQPVYTYDDEQJELIEDVDRGLYASENHCEILKCKLKLKA.FDV	958
Qy	962	LAHPONYFKYTQAEBSREMFRPSF-RLLASKVSRFLRIPYK	999
Db	959	LAHPONYFKYTQAEBSREMFRPSF-RLLASKVSRFLRIPYK	997

RESULT 11

ABP56758	ID	ABP56758 standard; Protein; 957 AA.
XX	DE	Mouse CEGP1 protein SEC ID NO: 6.
AC	KW	Human; secreted protein; epidermal growth factor subfamily;
XX	KW	epidermal growth factor; EGFR gene therapy; therapeutic; drug screening;
AC	KW	biological activity; immune response.
XX	XX	Mus musculus.
CS	XX	
XX	PN	
XX	PN	WO200210-08C-A2.
XX	PD	19-DEC-2002.
XX	PF	07-NAY-2002; 2002MC-US22278.
XX	PR	16-NAY-2001; 20CIUS-C855824.
XX	PA	{PEKE } PE CORP NY.
PA	PA	{GONG F.} GONG F.
PA	PA	{DFRAV/} DI FRANCESCA V.
XX	PI	Gong F., Di Francesco V., van C., Beasley EM:
XX	DR	NPI; 2003-1674C7//6.
XX	PT	New isolated human secreted peptides, useful for diagnosing or treating
XX	PT	a disease characterized by an absence of, inappropriate or unwanted
XX	PT	expression of the secreted protein and/or drug screening assays.

XX	PS Disclosure; Fig 2F-G; 185pp; English.	Qy	764 THR CIRCPVGYQPEF3KPNVSCPNTTDFDGTNTTCCKNRRGGELSDFTGYIESP 623
XX	The present sequence encodes a human secreted peptide (I) related to the epidermal growth factor (EGF) protein subfamily. (I) can be used in gene therapy, (I) and the nucleic acid molecules encoding (I) can be used as models for the development of human therapeutic targets, to aid in the identification of therapeutic proteins, and to serve as targets for the development of human therapeutic agents. (I) may be used in drug screening assays, in assays to determine the biological activity of the protein, to raise antibodies or to elicit another immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, or as markers for tissues in which the corresponding protein is preferentially expressed. (I) can also be used for diagnosing or treating a disease or disorder characterised by an absence of, inappropriate or unwanted expression of the protein. (I) is located to human chromosome 22. The present sequence represents mouse C2bp1 protein which is given in comparison with (I);	Db	722 THRCIRPLGTCPEFSKNNVSCPNTTDFDGTNTTCCKNRRGGELSDFTGYIESP 782
CC	Best Local Similarity 90.4%; Fred. No. 2.5e-298;	Qy	824 NYGNYPANTCTWTPPKRRLIVVPEIFLP2EDDCGDYJMRKTSSNSVTYETC 883
CC	Matches 862; Conservative 46; Mismatches 0; Gaps 0;	Db	782 NYGNYPANSETCTWTPPKRRLIVVPEIFPIEDCGDYLWRKTSSNSVTYETC 84:
CC	Query 44 EDVDECAQGLDDCHADALCQTPTSKCSCPKPGYQEGRCQDIDECNGEINGCQHDE: 103	Qy	854 QTYERPLAFTSRSKKKWIGFKSNEGSARGQVPPVYTYDQYELIEDVDRGRYASEN 943
CC	Db	2 EDVDECAQGLDDCHADALCQTPTSKCSCPKPGYQEGRCQDIDECNTLNGGCVHDCL 61	
CC	Query 104 NIPGNYRCTCFDGFMLAHDGHNCNLDECLENNGGCQHTECNGSYECCKEGFFLSDN 163	Qy	882 QTYERPLAFTSRSKKKWIGFKSNEGSARGQVPPVYTYDQYELIEDVDRGRYASEN 90:
CC	Db	62 NIPGNYRCTCFDGFMLAHDGHNCNLDECLENNGGCQHTECNGSYECCKEGFFLSDN 12:	
CC	Query 164 QHTCIHRSEEGLSCKMKDHGSHICKEAPRSVACECRPGFLAKNQDCLLTCHNGNG 223	Db	122 QHTCIHRSEEGLSCKMKDHGSHICKEAPRSVACECRPGFLAKNQDCLLTCHNGNG 181
CC	Db	224 CQHSCDTAQGECSCHPQHAKHTDSCLERDVLETEPNTSVVODKVRKLRY 283	
CC	Query 182 CQHSCDTAQGECSCHPQHAKHTDSCLERDVLETEPNTSVVODKVRKLRY 24:	Db	182 CQHSCDTAQGECSCHPQHAKHTDSCLERDVLETEPNTSVVODKVRKLRY 24:
CC	Query 284 ETCAVNRGGCDRTCKDTSTGTHCSCPVGFTLQLGKTCDDICOTRNGSCDHCFKNTV 343	Qy	242 ETCAVNRGGCDRTCKDTSTGTHCSCPVGFTLQLGKTCDDICOTRNGSCDHCFKNTV 343
CC	Db	344 SFDCGCKKGKFLLTDKEKSCQDVDECSLORTCSCINHPGTFACACNRGYLYGFTHCSD 403	
CC	Query 302 SFDCGCKKGKFLLTDKEKSCQDVDECSLORTCSCINHPGTFACACNRGYLYGFTHCSD 362	Db	404 TNECSINNGGQOQCINTVYECQCHPGKLUHNWKDCVKEWLPLPTSPPRSLHCK 462
CC	Db	362 TNECSINNGGQOQCINTVYECQCHPGKLUHNWKDCVKEWLPLPTSPPRSLHCK 42:	
CC	Query 464 SSGGDGFRLRCHSGHLSUDTITRSVTFKLNEGKSLKNAECEFGRPALPKHSSV 523	Qy	564 RLRKAIRTLRAVHREQFHQLSGMNLDVAKKPPRTSERQESCSVGHAEVSCPA 643
CC	Db	422 SSGGDGFRLRCHSGHLSUDTITRSVTFKLNEGKSLKNAECEFGRPALPKHSSV 48:	
CC	Query 524 KESFRYNTLTCSSGKQVPGAPGRSPKTFMPTFELETNQKEYTASCDLSCTVKRTK 582	Db	542 RLRKAIRTLRAHREQFHQLSGMNLDMAKTPSRVSGOHEETCGVGOGEESQCVSCPA 60:
CC	Db	482 KESFQYANLTCSPKQVGAQRLNAPKEMTFVEBRTVEKEYTASCNLSCTVKRTK 54:	
CC	Query 584 RLRKAIRTLRAVHREQFHQLSGMNLDVAKKPPRTSERQESCSVGHAEVSCPA 643	Db	644 GIYDGARERCILCPNGTQNQEGMTCEP CPPRGNNSGAJKTPEAMNSECQGQGPGEY 703
CC	Db	542 RLRKAIRTLRAHREQFHQLSGMNLDMAKTPSRVSGOHEETCGVGOGEESQCVSCPA 661	
CC	Query 602 GIYDGARERCILCPNGTQNQEGMTCEP CPPRGNNSGAJKTPEAMNSECQGQGPGEY 703	Db	704 SADGFAPCQLCALGTFOPEAGRTSCFPGGGLATKHQGATSFQDCETRVOCSPGHFYNT 763
CC	Db	662 SANGFAPCQLCALGTFOPDVGRISLUSCGGLPLTHUGATSQDCETRVOCSPGHFYNT 721	

Sequence 988 AA;

XX

ABP5675 standard; protein; 988 AA.

XX

ABP5675;

AC

XX

DT 31-MAR-2003 (first entry)

DE Human epidermal growth factor related secreted protein SEQ ID NO:2.

XX

KW Human; secreted protein; epidermal growth factor protein; subfamily; epidermal growth factor; EGFP; gene therapy; therapeutic; drug screening; bioactivity; immune response; chromosome 22.

XX

Homo sapiens.

GS XX

W020021C080-A2.

PN XX

PD 19-DEC-2002.

XX

PF 07-MAY-2002; 2000WD-13322258.

PR XX

PA 16-YAY-2001; 2001US-0855624.

PA :DEKE : PE CORP NY.

PA :CONG / :GCNG F.

PA :DFRA / : DE FRANCESCO V.

XX

Gong F, Di Francesco V, Yan C, Beasley EM;

PI XX

DR 2293-167407/-6.

N-PSDB: AB222653, AB222654.

PR New isolated human secreted peptide, useful for diagnosing or treating a disease characterized by an absence of, inappropriate or unwanted expression of the secreted protein, and in drug screening assays -

XX

PA

PS Claim 1; Fig 2A; 185pp; English.

CC The present sequence represents a human secreted peptide (I) related to the epidermal growth factor (EGF) protein subfamily. (I) can be used in gene therapy, (I) and the nucleic acid molecules encoding (I) can be used as models for the development of human therapeutic targets, to aid in the identification of therapeutic proteins, and to serve as targets for the development of human therapeutic agents. (I) may be used in drug screening assays, in assays to determine the biological activity of the protein, to raise antibodies or to elicit another immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, or as markers for tissues in which the corresponding protein is preferentially expressed. (I) can also be used for diagnosing or treating a disease or disorder characterised by an absence of, inappropriate or unwanted expression of the protein. (I) is located to human chromosome 22.

XX

Sequence 988 AA;

disease, Parkinson's disorder, immune disorders, haemopoietic disorders, and various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X, wasting disorders associated with chronic diseases, and cancer. The isolated protein, its encoding polynucleotide or an antibody created from the protein are useful for treating or preventing neurological disorders such as epilepsy, stroke, mental disorders including mood, anxiety, schizotypal disorders, disorders of vesicular transport, such as cystic fibrosis, diabetes mellitus, goiter, gastrointestinal disorders including ulcerative colitis, other conditions associated with abnormal vesicle trafficking including AIDS, allergic reactions, multiple sclerosis and rheumatoid arthritis. A cell comprising the vector of the invention is useful for producing non-human transgenic animals. The polynucleotide of the invention can be used to treat disorders by gene therapy. This sequence represents one of the isolated NOX proteins of the invention.

Sequence 988 AA;

Query Match: 65-31; Score: 3662.5; DB: 23; Length: 988;
Best Local Similarity: 62.2%; Pred. No.: 3e-219; Indels: 119; Gaps: 14;
Matches: 652; Conservative: 113; Mismatches: 164;

Qy 10 GAA--WAVLILLLPPLILLAGVAPPGRGAA----PQDVDEAQGLDDCHADAL 61
Db 2 GRAAVRWHLCVLLA-----GTRGLLGGSLPGSVDVECESGDCHDAI 49
Qy 62 CQNTPTSYKCSKTPGQGEHQGQEDIDECKE-LNGGYHDVCDINTGPNYRCTFDGPNLA 120
Dg 50 QNAPSKSYKCLPKPGKKGHQCGDIDECKENDYNGGCYHECINPGNYRCDFGPMLA 109
Qy 121 HDGHNCNLDDVDECLENNGGCQHTCYNMGSYCCCKEGFPSSDNOHTCHRSBEGLSCKNK 180
Db 110 HDGHNCNLDDVDEQDNNGGCQHQICDNNMGSYCCCKEGFPSSDNOHTCHRSBEGLSCKNK 169
Qy 181 DHGCSHICKAEPRGSVACECRPFELAKNORDCILTCNHCNGGCOHSCTDTADGPECSCH 240
Db 170 DHGCAHICRETKGGYACDPRGFDAQNNDCTCNCNNGGCOHSCTDTDGPTECCH 229
Qy 241 PQYKMHITGCSRSLERETVLEUTESNTTSVNDG3KRVKARLIMETCAANGGDRKNT 305
Db 230 QKYPADHGRTCI-----ETCAVNNNGGCCRDKD 259
Qy 301 STGVKSSCPVGFTLQLDGKTCKDIDECQTNNGGDFHCKNIVGSPDCGCKGPKJLDEK 360
Db 260 ATGVRCSCPVGFTLQLDGKTCKDIDECQTNNGGDFHCKNIVGSPDCGCKGPKJLDEK 319
Qy 361 SCQDVDECSLDTCDSNCINERFGTACACRGTYLGLFGTGTNECSINNGGCCYCVN 420
Db 320 TCDIDECSFERTCDHCINCSPGSFQLCLRGYIYGTHGVDCEXSNSCDQCVN 379
Qy 421 TGSYBQCHPGYKUHNWKDCVEV-KGLLP-SPSVPVSLHCGKSGGDBCFRC-----
Db 380 TGSYBQCHPGYKUHNWKDCVEV-KGLLP-SPSVPVSLHCGKSGGDBCFRC-----
Qy 475 -----HSGIHLIS---SDVTIRTSTVTPKLNENGKSL 502
Db 440 FVPDSENSVYLSGVGPGQKALQRKNGTSGLGRSPDAPTRIKARFKRDARKHL 499
Db 500 R-----PHSQRARAKETAQPLDHCFTYFVTKCDSSKK-RRRRKPSK 544
Qy 503 KNAELFFEGLRPALPEKHSSVKEFSEFR-----YVNLTCSGGKQYPGAPGRSPPK 552
Db 553 M-FITVFELETINQKEYTASCDLSCIVKREKLRKA1STJRKAHVREOFHLSGMNJD 612
Qy 545 VSHITAEEFIEYKMEASDTCEDALCLRKRAEQLAAIKTLRKIGRQOFYVQVSCTEYE 604
Db 612 VAKKPPRTSERAESCCVGGHAENQCVSRAFTYDGAERCLCPNNTFONEBEGOMTC 671
Qy 605 VAQRPAATALEGQG-ACGAGQVLQDSKSCVAGCPGTHGGELGQCPMC6TYQDMEGQLSC 663
Db 672 EPCPRPGNSGALKTPEAWNNSCSEGGLCPQGEYSADGFAPCOLCALGTFQEPAGRTSCEPC 731
Qy 664 TPCP---SSDGLGLPGARNVSECGQGSPFSADEGFKPQACPVGTVQEPGRTGCEPC 720

Qy	732 SGGLATKNGGATSFQDCEPVOCSRGHNTTTHRCRCPVHTYQEEGRNCGCSCPRT 731
Db	721 GGGLTREGSTSFCDEAKVHCSPEGHNTTTHRCRCPVHTYQEEGRNCGCSCPRT 780
Qy	792 TTDGDSNTTQKNRGGEGDFGTYESPNGFNYPANTECWNTNPPIPRILTV 851
Db	761 STDGDSNTTQKNRGGEGDFGTYESPNGFNYPANTECWNTNPPIPRILTV 840
Qy	852 PEIPFPIEDCGYLWNRKTSSNSVTTYETCQYERPIAFTSRKWLQFKSNEGNSA 911
Db	841 PEIPFPIEDCGYLWNRKSAPSITTYETCQYERPIAFTSRKWLQFKSNEGNSA 936
Qy	912 RGFCVPPVTDYDQYCEOLEDIVDGRGLYASENHCETLKDKKLKIAFCVLAHPQNYFKYT 971
Db	921 KGFOVPPVTDYDQYCLDITDGRGLYASENHCETLKDKKLKIAFDVLAHPQNYFKYT 965
Qy	972 AQESREMFFERSFLRLASKSVRSLPRLPYK 999
Db	961 AQESREMFFERSFLRLRSVSLPRLPYK 988
Qy	RESJLT 14 AAU73172
Db	AAU79-72 standard; Protein: 1C06 AA. XX
Qy	ACAU79-72; XX
Db	02-JUL-2002 (first entry) XX
Qy	Human MEGF/FIBRILLIN-like protein NCv8. XX
Db	Human: MEGF/FIBRILLIN-like protein; receptor: NOCV, NOV8; XW vaccine; cancer: tumour; bone disorder; avascular necrosis; artery; XW haemato poetic disorder; immune disorder; endometriosis; renal disease; XW infection; inflammatory disease; lung disease; scleroderma; ataxia; XW bowel disease; appendicitis; blood disorder; cardiovascular disorder; XW graft versus host disease; GVHD; lymphoedema; brain disorder; XW ocular disorder; hepatitis C virus infection; cardiac disorder; XW autosomal dominant deafness; DFNA-2; chromosome 6. XX Homo sapiens. XX NC200214368-A2. XX
Qy	21-FEB-2002. XX
Db	16-AUG-2001; 20G1WC-US25624. XX
Qy	16-AUG-2001; 2000US-225692P. XX
Db	16-AUG-2000; 2000US-225693P. PR
Qy	16-AUG-2000; 2000US-225817P. PR
Db	18-AUG-2000; 2000US-226236P. PR
Qy	22-AUG-2000; 2000US-227385P. PR
Db	23-AUG-2000; 2000US-227395P. PR
Qy	24-AUG-2000; 2000US-227492P. PR
Db	14-XR-2001; 20031S-2275952P. XX
Qy	{CURA}: CURAGEN CORP. XX
Db	Zerhusen BD, Padigaru M, Spytek RA, Spaderna SK, Ganguli EA; Rastelli L, Burgess CB, MacLunder K, Snimkets R, Mishra V; Verrier CAM, Szekeres ES, Grossie KW, Alsobrook JP, Liu X; Gerlach VL, Elizetman K, Smithson G, Smithson J, Peyman J, Macdougal J; XX
Qy	WPI: 2002-323571/36. DR
Db	N-PSDB; ABK4392. XX
Qy	Novel cytoplasmic, nuclear membrane bound and secreted NOX peptides, useful for treating cancers and tumours, bone disorders.
Db	PT

Pt's disease, haemopoietic disorders, spinal diseases and immune disorders -
 PT
 XX
 PS Claim 1; Page 77; 234pp; English.
 CC The present invention relates to new isolated NOVX polypeptides named NOV1-NOV9. The invention can be used for identifying an agent (a cellular receptor or downstream effector) that binds to the polypeptide. The molecules of the invention are useful for treating or preventing NOVX-associated disorders in humans. The antibody of the invention is useful for determining the presence or amount of NOVX in a sample, and for treating a pathological state in a mammal. The method of the invention is useful for determining the presence of an amount of NOVX in a sample which is used as a marker for cancerous cell or tissue type. The molecules of the invention are useful in the manufacture of a medicament for treating or preventing cancer, tumour, bone disorders, atherosclerosis, allergy, haemopoietic disorders, immune disorders, endometriosis, renal diseases, infections, inflammatory diseases, lung diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood disorders, cardiovascular disorders, graft versus host disease (GVHD), lymphoedema, brain disorders, ocular disorders, hepatitis C virus infection, cardiac disorders and autosomal dominant deafness (DFNA-2). The present amino acid sequence represents the human MEGF/FIBRILLIN-like protein NOV8 that is one of the NOVX proteins described in the invention. This sequence is encoded by the human MEGF/FIBRILLIN-like protein NOV8 gene located on chromosome 6.

XX Sequence 1006 AA;
 Query Match 61.4%; Score 3443.5; DB 23; Length 1006;
 Best Local Similarity 59.6%; Prey: No. 1.e-205;
 Matches 601; Conservative 150; Indels 41; Gaps 9;

Qy 4 AGRNRPGAAAVLVLILLPLLLAGAAPPGRGRRAGPQDVDECAQGLDDCHADALCQ 63
 Db 26 AAQCKEGGARGLGAVLAAPQGL--GAPEGGGRCMNPLDVDECVEGSTNTIDAICQ 92
 Qy 64 NPTPSYKCSCKPKGYQGERGRQCDIDECGNELNLNGCVH3CLN1-PGNYTRCTCFDGFMLAHDG 123
 Db 83 NPTPSYKCSCKPKGYQGERGRQCDIDECGNELNLNGCVH3CLN1-PGNYTRCTCFDGFMLAHDG 142
 Qy 124 HNCLDVECLENNGCQHTCUNVNGSYECCRGKGFPLSDNQHTCUNHSEEGLSCKMKDHG 183
 Db 143 HNCLDVECLEBNGCQHSVNMGSYBHCRCBFFELDSQNQHTCUNQPEBGNCKAKQHG 202
 Qy 184 CSHTICKEAPGSVAECCRPFELAKNQDCILTCUNHNGNGCOHSDDTADGPECSCHPOY 243
 Db 203 CAHICRETPKGIAECCRPFELTKNQDKLUTCNNGGCOHTCDEQGRGCHHKF 262
 Qy 244 KMFITGRSCLSEREDTVELETTSNTTSDGDKRVRKL---LMETCAVNINGCORTCK 298
 Db 263 VLHTDGKTCI-----GERLLEQH-PTQAVSNETCAVNGGDSKCH 303
 Qy 299 DTSTGVHCTSPVGFTLQLDQKTKDIDSGQTTRNGCDHFCNIVGSDGCKRKGFKLTC 358
 Db 304 DAATGVHCTCPVGMQLPDRKTCRDIDECRLNNGCDHICRATVGSFECCKGKYKJLN 363
 Qy 359 EKSCGDVDECSLDRTCDHSCINHFGTFACANCGTYLYGFTHGDTNECS1-NNGGCCQVC 418
 Db 364 ERNCODIDEESFDRCDHCTCVENTGFSQCLHRYLIGTHGVDDECNSNRRGRFGC 423
 Qy 419 VNTVGSYEECCCHPGEY-KLJWWNKXDCVENVKGLLTTSVSPRVS1-HCGKSGGDCFLCHSG 477
 Db 424 INTPGYQCTPAGQGRHLNGKDCTEP1.RCQGSPGASKAVL1QNSRKC1KTCALTGST 483
 Qy 478 IHSSDDVTTRTSVTFKLNGKCSLKNAEFLPESL---RPALPEKISSKRSFRVNL1 533
 Db 484 IPLEAVLSSLKORAFKIKDAKCRLAGSENFGDVPGVDFKPSLPGGAPCSCECOVTE.HLK 543
 Qy 534 CSSGKQVPGAPGPSPTPKEMNITYVEPELETOENQEVETASCDLS1CIVKTERLRKA1.PTLR 593
 Db 544 CDSSRKKGKGRARRPSPCSQQFLLP---DTGHGPPASCGPLCLRQMRERLKGSLKMLR 599

Cy 534 KAVHREQFHQLQSGYVJLVAKKPPTSERQA---ESCGVGGGHAENCYSCRAGTYYDGA 650
 Ds 650 K5-NQDRPLIRLRLAGLQDYLAKPLVAGERAEPMECRGHRADSPSvvcPCGTVYHGQ 659
 Cy 651 REPCILCNGTQNERQMTCEPPCPGPNSGALKTPEANMSECGGLCQGEYSADGFAPI 710
 Db 660 TEQCYPCFAAGTQPEREQLSCLCPREKDQSPKPKHAGPKCQR-CRRKHSSDGFKP 718
 Qy 711 COLCALGTQPEAGRITSFPCGGGLATHKGATSFCCTCETRVCQS2GHTNNTTHRCC 770
 Db 719 CQPQPRGYQPEAGRITCPCCGGLT-KHEA1SFCD2TQVQCS2GHTNNTTHRCC 778
 Cy 771 PVGTYQPEFGKONCVSPGNNTQDGSTNTIOTCKDRCRGCGELGDTGYTESPNPGNYP 830
 Db 779 AMGYYCQPERQNCSPSRGNTTDFDSTSAAQKARQCGEL-GETGYESPNGNYP 938
 Cy 831 ANPECTATINPPPKRLLINPBIPELPIEDDGGY-AWAKTSSSSNTVTTETCCYERPI 890
 Ds 839 AGVECINNINPPPKRLLINPBIPELPIEDDGGY-AWAKTSSSSNTVTTETCCYERPI 898
 Cy 891 APTRSKZLWIOFKSNEGNSANGFQPYVTDYDQE1-EPD1VRCRLLYSENHEILKD 950
 Db 999 APTRSRSLWINKFTSANSANGFC1PYVTDYDQE1VLCEDIVRGRYASENHEILKD 958
 Qy 951 KK-IKTA-PDVLAPQNTFYKTAQESRFNPSSFIRLRSVSRFPRYK 999
 Db 959 KKLIKAFPEVLAHPQNYFKYT-EKHSMPLKSFIKLRSKVSFLPRYK 1006

RESULT 15
 ABU12093
 ABU12093 standard; protein; 974 AA.
 AC ABU12093;
 XX DT 17-FEB-2003 (first entry)
 DE Novel: Human epidermal growth factor-like protein #3.
 XX KW Gamma-aminobutyric acid receptor-like protein; depression; stroke;
 KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;
 KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
 KW Alzheimer's disease; alcoholism; vigilance; muscle tension;
 KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
 KW arrhythmic right ventricular dysrhythmia; renal disease; diabetes;
 KW Epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
 KW haematopoietic stem and progenitor cell like protein; cirrhosis;
 KW sultor transferase-like protein; cholangitis; hepatitis; hepatitis; diabesity;
 KW developmental disorder; Syntaxin-like protein; myoid liposarcoma;
 KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloid leukaemia;
 KW transgenic animal 1.
 XX CS Home sapiens.
 XX PN US2002123612-A1.
 XX PD 25-SEP-2002.
 XX PR C3-JUL-2001; 2001US-C8998570.
 PR 19-APR-2000; 2000US-1982932.
 PR 20-APR-2000; 2000US-198645P.
 PR 25-APR-2000; 2000US-199476P.
 PR 26-Apr-2000; 2000US-199880P.
 PR 26-Apr-2000; 2000US-200024P.
 PR 09-JUN-2000; 2000US-20052P.
 PR 03-JUL-2000; 2000US-210809P.
 PR 17-JUL-2000; 2000US-215855P.
 PR 11-AUG-2000; 2000US-228591P.
 XX PA {GERLACH V. L.; ELLERMAN K.; ELLELSE; GERLACH V. L.; ELLERMAN K.;

PA	(MACD/) MACDOUGALL J R. (SMIT/) SMITHSON Q.	DB	287 CRATCKED DECLININGCDH CRNIVGFFCSCRGKYLTTNEFNQCLBECGFTRCFH 346
XX		QY	377 SCINHPCFTAACNRGTYLTGFTHCGDTNNECSINTNGCCQCYCORTVGSYECQCHPGY KL 435
P1	Gerlach VL, Ellerman K, Macdougall JR, Smithson G;	DB	347 ICNTNPSSFC-CCHRSLVYKTHGDDCISNGGFC-NTPGYYCCTPAGCQL 406
XX		QY	436 HANKKOCYEVNSJLPTSVSPAVS1HNGKSEGGDCFLRCHSGIHL-SSEUVTIRTSVTF 493
DR	WPI: 2003-06615/06.	DC	407 HNGKCCTEPJKCCGSPASAMLSNCNRSEKKDCAUTPSRARFLPZAVLSIKRASF 466
XX	Novel polypeptides and nucleic acids which are members of epidermal growth factor, complement receptor families for diagnosing and treating psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's disease -	QY	454 KUNEGKC-SILKNAELFPEGLRPALPEKISSKVESFRYVNLTSSEGGKQPGAGRSPSTPK 551
PT	Claim 1: Page 23-24: 91PP: English.	DB	457 K1KDAKCR-LHANKGKTEAERTCSEGGAPECCVTFHLKCDSSRKGRBARTPPGK 526
PT	the invention describes an isolated POLYX (POLY1-1), polypeptide and its variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the polyribonucleotides encoding them (I) and an anti-POLYX-antibody (III) are useful for treating or preventing a pathology associated with human POLYX disease. POLYX polypeptide is also useful for identifying an agent that binds to POLYX and a cell expressing POLYX is useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. (III) is useful for treating a pathological state in a mammal and for determining the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like proteins) are useful for the treatment of psychiatric and medical conditions, depression, stroke, Parkinson's disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension, epileptoic activity and memory functions, cardiomopathy and arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth factor like proteins) may be useful for treating cancer, aberrant angiogenesis, renal disease and diabetes. POLY12 (haemopoietic stem and progenitor cell like protein) may be useful for treatment of leukaemia, lupus and anaemia. POLY13 (sulfotransferase-like protein) may be useful for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism and developmental disorders. POLY14-16 (Synaxin-like proteins) may be useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid liposarcoma and acute myeloid leukemia, and POLY18 may be useful in treatment of cancers. Cells comprising (I) are useful for producing non-human transgenic animals which are useful for studying the function and/or activity of POLYX protein and for identifying and/or evaluating modulators of POLYX protein activity. This is the amino acid sequence of a novel human protein.	QY	552 ENP-ITVEFELETFENQKEVTAACOLLSIVTRERLARKA-RTLKVAVRQPHIQLSGMNL 610
PT	Sequence 974 AA;	DC	527 EYTRTLELEAVRAEETTASCGLPCLRQYERRLKGSLKMJKRSINSQCBFLRLLAGLDY 586
PT		QY	611 DYAKKPRTSERQA--ESCGVGQGHAEINOCVSRACTYDGAERCI:CPNGTFCNNEG 667
XX		DB	587 ELAHKPGLVAGRAEPMESCRPGQERAGTCVSCQFQGTIVHGQECPAGTCFQERG 646
PS		QY	668 QMTCEPCPRPGNSGALKTPEANWNSCCG:CGPGBYSADGAPCOLCALGTPEEAQGTS 727
XX		DB	647 QSCDCLPGSDARGPL--GATINVTCAGCOPGHSQDFKQPQCPBRSTYQEAQRTL 733
CC		QY	728 CFPCCGGIATKHOGATSFGCCTTRYCQSCPGHFYNTT:HCIRCEVGTQDPFGKNCYC 787
CC		DB	704 CFPCCGGATTTKHEGA,SFQGQDQTKYQCPGAYNTSIHCICRANGSQDFEQNFCSR 763
CC		QY	758 PGNTT:PDGSTNTTICKNRRCCGEG:GDFGTYISPNPNYGNYFANTECWT:IPPPRRL 847
CC		DB	764 PGNTT:PDGSTTSVAQCNRACGGELGEFTSYIBPNYGNYPAGEVCWNTPPKKI 823
CC		QY	848 LIVPEPLPDEDGSDYLYWKT:SSNSNTTYETCQYERPIATTSRSKKLMCFKSNE 907
CC		DB	824 LIVPEPLPSEDDECSDVLMRNKSSPS:TTYETCQYERZ:AAFTARSKLWINEFTSE 883
CC		QY	903 CNSARGFQPVYTYDYECE:ZEDIVDRGRLYASENHCSBILKCKKLIZL:FCVLAHPNY 967
CC		DB	864 ANSARGQ:PVTYDYEQIVEDVRDGLYASENHQETLKDXKL:KAFFEVAJAHPSNY 943
CC		QY	368 FKYTAGCSEREMPRFSEFLRSKVSRLFRPYK 999
CC		DC	944 FKY-T-EHKEM-JPKSFKLRSKVSFLFRPYK 974
SC			
Query Match	60.9%	Score	3415.5;
Best Local Similarity	60.1%	DB	24; Length: 974;
Matches	596;	Conservative	140; Pres. No. 6.e-204;
		Mismatches	219; Index:
			37; Gaps 9;
QY	22 LPPJLLLAGVPPGRGAA3PQEVDEAQGLDDCHADLONTPTSYCKCCKEGYQGEG 8:		
	6 VPGCLVILVNHARAQVSKAAQDVDECEVGTNCHIDAIQNTPTSYCKCCKGTYGDS 65		
DB	82 RQCEDIDECCNEELANGCVHDCLN1PGNYTRCTCPDGFMLAHDHNCNDVDEAEGNGGCQH 41		
QY	66 KHCKDVODECREDRAGCYHDCVNIPGNVYRCTCDGFLAHDHNCNDVDEAEGNGGCQ 125		
DB	142 TCVNMGSYECCKKEGFFLSDNOHTCTHRSSEGLSMKNDQHCSHCKEAPRGSYACECR 20-		
DB	126 SCVNMGSYEHCRCGPFLSDQETCOPPEGRNMRHGCANHCRETPGGFACECR 185		
QY	202 PGFELAKNORDCILTNHNGNGCQHSCDDTADGPECSCHPPQKHTDGRSSLERETVLE 261		
DB	186 PGFELAKNQDCK:TCNYGNNGCQHCTDDEQSPRCGHKIVFLHTDGKTI----- 237		
QY	262 VTESHTSVUDGDKRVKRL---LMETCAVNGGCDRTCKDTSGYHCCCPVGFTLQ 316		
DB	238 -----GERMLEQHPTQAVSNETCAVNGGCDSKHDAATSVHCTCPVGFMLQP 286		
QY	317 DGKTKDIDCQTRNGGCDHFCRNIVGSPDGFGRKFKLTDKSKCQDVDSLRTCDH 376		

Search completed: October 22, 2003, 21:27:32
Job time : 34 secs

Result	NC	Score	Query Match	Length	DB ID	Description
1	5605	100.0	999	9	US-09-747-371-2	Sequence 2, App1;
2	5605	100.0	999	15	US-10-177-293-45	Sequence 45, App1;
3	5581	99.6	997	15	US-10-176-847-50	Sequence 5C, App1;
4	5428	96.8	964	12	US-09-555-324-5	Sequence 5, App1;
5	5056	90.2	997	9	US-09-747-371-3	Sequence 3, App1;
6	4935	88.0	957	12	US-09-855-834-6	Sequence 6, App1;
7	3675	65.6	988	12	US-09-855-834-6	Sequence 2, App1;
8	3675.5	65.6	988	15	US-10-195-442-2	Sequence 2, App1;
9	3412.5	61.0	974	10	US-09-898-510-14	Sequence 14, App1;
10	3412.5	61.0	974	11	US-09-839-446-14	Sequence 14, App1;
11	3397	60.6	993	12	US-10-239-663-36	Sequence 36, App1;
12	3394	60.6	1009	10	US-09-898-510-16	Sequence 16, App1;
13	3394	60.6	1009	11	US-09-839-446-16	Sequence 16, App1;
14	3290.5	58.7	964	12	US-10-239-663-58	Sequence 51, App1;
15	3290.5	58.7	965	12	US-10-221-097-51	Sequence 51, App1;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Match	Score	DB	ID	Description
1	100.0%	DB	9	Length 399;
2	Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;
3	Matches 399;	Conservative	0;	Gaps 3;
4	1	MGVAGRNPFGRGAWAVLILLLPPUULLAGVPPSRGRAGPCEVOECAGLDCHADA 60		
5	1	LCQNTPTSYKCSCKPKGYCGESRQCEDIDECNENLNGCYHDLNPNSYRCTCFDGFMIA 120		
6	1	MGVAGRNPFGRGAWAVLILLLPPUULLAGVPPSRGRAGPCEVOECAGLDCHADA 60		
7	1	LCQNTPTSYKCSCKPKGYCGESRQCEDIDECNENLNGCYHDLNPNSYRCTCFDGFMIA 120		
8	1	MGVAGRNPFGRGAWAVLILLLPPUULLAGVPPSRGRAGPCEVOECAGLDCHADA 60		
9	1	LCQNTPTSYKCSCKPKGYCGESRQCEDIDECNENLNGCYHDLNPNSYRCTCFDGFMIA 120		
10	1	MGVAGRNPFGRGAWAVLILLLPPUULLAGVPPSRGRAGPCEVOECAGLDCHADA 60		
11	1	LCQNTPTSYKCSCKPKGYCGESRQCEDIDECNENLNGCYHDLNPNSYRCTCFDGFMIA 120		
12	1	HDGHNLCLDVECLENNINGCQHTCCTVNYGSYCCKEGFLSNCQHTCIRSEEGLSCMNK 180		
13	1	HDGHNLCLDVECLENNINGCQHTCCTVNYGSYCCKEGFLSNCQHTCIRSEEGLSCMNK 180		
14	1	HDGHNLCLDVECLENNINGCQHTCCTVNYGSYCCKEGFLSNCQHTCIRSEEGLSCMNK 180		
15	1	HDGHNLCLDVECLENNINGCQHTCCTVNYGSYCCKEGFLSNCQHTCIRSEEGLSCMNK 180		

Y	b	181	DHGCHICKEAPRGSSVACECRPGFELAKNORDCILTCNHNNGCQHSDDTAGGBCSCH	240
Y	b	181	DHGCHICKEAPRGSSVACECRPGFEEAKNORDCILTCNHNNGCQHSDDTAGGBCSCH	240
Y	b	241	PQYKMHDTGRSCLEREDTVLEVTESTNTTSVVDGCKRKVRKRRLMETCAVNNGGCDRTCKDT	300
Y	b	241	PQYKMHDTGRSCLEREDTVLEVTESTNTTSVVDGCKRKVRKRRLMETCAVNNGGCDRTCKDT	300
Y	b	301	STGVHCSCPVGFITLQDGKTKDIDECQTRNGGCDHFCKRN:VGSFDGCKKGFKLTDEK	360
Y	b	301	STGVHCSCPVGFITLQDGKTKDIDECQTRNGGCDHFCKRN:VGSFDGCKKGFKLTDEK	360
Y	b	361	SCQDVDECSLDRTCHSCINHPGTACACNRGYTLIGFT:EGDNTECSINGGCCQCVN	420
Y	b	361	SCQDVDECSLDRTCHSCINHPGTACACNRGYTLIGFT:EGDNTECSINGGCCQCVN	420
Y	c	421	TVGSYECQCHPGYKLHWNKDKCVEYKGLPTSVSPRVSLHCGKSGGDGFLRCSGHL:	480
Y	c	421	TVGSEQCQHCGYKLHWNKDKCVEYKGLPTSVSPRVSLHCGKSGGDGFLRCSGHL:	480
Y	b	481	SSDVTTIRTSVTFKLNEGKCSLKNAELFPECLRPALPEKSSVKESSFRYVNLTCSGKQV	540
Y	b	481	SSDVTTIRTSVTFKLNEGKCSLKNAELFPECLRPALPEKSSVKESSFRYVNLTCSGKQV	540
Y	b	541	PGAPGRPSTKREMFITVEFELETNQEVTAASCDSL:IVRTERKLRKATRLRKAVHREQ	600
Y	b	541	PGAPGRPSTKREMFITVEFELETNQEVTAASCDSL:IVRTERKLRKATRLRKAVHREQ	600
Y	b	601	FHQSGMNLDVAKKPRTRSERQASCVGCGHAEENGCVSCRAGTYDYGDRARERCCLCPNG	660
Y	b	601	FHQSGMNLDVAKKPRTRSERQASCVGCGHAEENGCVSCRAGTYDYGDRARERCCLCPNG	660
Y	b	661	TPQNEGOMTCEPCPRPGNSGALKPPEAMNMSECGLCQGEYSADGPACQJCLG:FAQ	720
Y	b	661	TPQNEGOMTCEPCPRPGNSGALKPPEAMNMSECGLCQGEYSADGPACQJCLG:FAQ	720
Y	c	721	PEAGTSFCPGGGLATHGOGATSDQCETRVQSPGHFYNTNTTHRCIRCPVGTYQPEFG	780
Y	c	721	PEAGTSFCPGGGLATHGOGATSDQCETRVQSSGHFNTNTTHRCIRCPVGTYQPEFG	780
Y	c	781	KNNCYSCPNTMTTDGDSTNITOCKMRRCGELGFTGYIESPNYPGNYPANTECTWT:N	840
Y	c	781	KNNCYSCPNTMTTDGDSTNITOCKMRRCGELGFTGYIESPNYPGNYPANTECTWT:N	840
Y	c	841	PPPKRILIVVPEIPLIEDDGDYLVMRKTSSNSVTTYETCQTYERPLAFTSSRSKLN	900
Y	c	841	PPPKRILIVVPEIPLIEDDGDYLVMRKTSSNSVTTYETCQTYERPLAFTSSRSKLN	900
Y	c	901	IQFKSNEGNSARGFQPVYTYQEDQELIEDIVRGRYLASENHQIILDKKJIKAI,FDV	960
Y	c	901	IQFKSNEGNSARGFQPVYTYQEDQELIEDIVRGRYLASENHQIILDKKJIKAI,FDV	960
Y	c	961	LAHPONYFKTAQESREMFPRTLLRSKVSRELRYK	999
Y	c	961	LAHPONYFKTAQESREMFPRTLLRSKVSRELRYK	999

RESULT 2
5-10-177-293-45
Sequence 45, Application US10177293
Publication No. US2003C12412BA1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarapu, Marjula
APPLICANT: Kamatkar, Siubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhe;
APPLICANT: Yu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John

APPLICANT: Meyers, Rachelle B.
APPLICANT: Bast, C., Robert C.
APPLICANT: Hotchbagyi, Gabriele N.
APPLICANT: Fuszta, Latos
APPLICANT: Meric, Funda
APPLICANT: Shahn, Alyseguin
APPLICANT: Milis, Gabor B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-036
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/1299,887
PRIOR FILING DATE: 2001-06-11
PRIOR APPLICATION NUMBER: US 60/131,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/106,501
PRIOR FILING DATE: 2001-07-21
PRIOR APPLICATION NUMBER: US 60/1325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/1362,585
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/1xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 306
SEQUENCE SOURCE: FastSEQ for Windows Version 4.0
SEQ ID NO: 45
LENGTH: 999
TYPE: PRT
ORGANISM: Homo sapiens
DBSOURCE: NCBI
DBID: 10-77-293-45

Db	541 PAGPRSPTPKMFITFELTNOKEVTTASCDLSCIVARTERKJRKATNLRAVHREC	600	Db	239 FQYKMTEDGRSCLEREDTVLETESENNTSVNDGDKVKGRLMETCAVNNGCDRTCKDT	298
Qy	601 FHQLSGMNLDAKKPRTSERQAEESCVGQGCHAENQCVSSRAGTYD GARERCILCPNG	660	Qy	321 STGVHCSRGPERLQLDGKTKQIDEQTRNGCDBHFCKNVGSFDGCKKGFKLTDK	360
Db	601 FHQLSGMNLDAKKPRTSERQAEESCVGQGCHAENQCVSSRAGTYD GARERCILCPNG	660	Db	299 STGVHCSRGPERLQLDGKTKQIDEQTRNGCDBHFCKNVGSFDGCKKGFKLTDK	358
Qy	661 TPENEEGOMTCBPCPRGNSGALKTPBANNSECGGLCQPBEYSADGFAPQLCALGTQ	720	Qy	361 SCQDVDESSLQTCDCPNSGALKTPBANNSECGGLCQPBEYSADGFAPQLCALGTQ	420
Db	661 TPENEEGOMTCBPCPRGNSGALKTPBANNSECGGLCQPBEYSADGFAPQLCALGTQ	720	Db	359 SCQDVDESSLQTCDCPNSGALKTPBANNSECGGLCQPBEYSADGFAPQLCALGTQ	418
Qy	721 PEARITSCFPCCGGLATHQGATSFDQDCTRVQCSPGHFTNTHRCIRCPVGTYQPEFG	780	Qy	421 TWSYECCCHGPGYKLHNKKDCDVEVKGLPLTSVSPRSVLHGKSGSGDGCPLRCHSGIHL	480
Db	721 PEARITSCFPCCGGLATHQGATSFDQDCTRVQCSPGHFTNTHRCIRCPVGTYQPEFG	780	Db	419 TWSYECCCHGPGYKLHNKKDCVKGJLJPLTSVSPRSVLHGKSGSGDGCPLRCHSGIHL	478
Qy	781 KRNVCSCPNTTDFDGSTNTITOCKNRCCGFLGDFGTYTESPNYPGNYPANTECTWTIN	840	Qy	481 SSZVTRTTSVFKJNEGKCSJKAELFPEGGLPAPLPERKSSVKSFRYVNLTCSSGKCV	540
Db	781 KRNVCSCPNTTDFDGSTNTITOCKNRCCGFLGDFGTYTESPNYPGNYPANTECTWTIN	840	Db	479 SSZVTRTTSVFKJNEGKCSJKAELFPEGGLPAPLPERKSSVKSFRYVNLTCSSGKCV	538
Qy	841 PPKRRLLIVPEIFPLEDDGDLVMRKTSNSNTYTETCQTYERPIAFTSRSKLW	900	Qy	541 PGAPGRPTPKEMFITFELTNCKEVTAASCDSCLSIKVTRKAVHREQ	600
Db	841 PPKRRLLIVPEIFPLEDDGDLVMRKTSNSNTYTETCQTYERPIAFTSRSKLW	900	Db	539 PGAPGRPTPKEMFITFELTNCKEVTAASCDSCLSIKVTRKAVHREQ	598
Qy	901 IQFKSNEBNNSARGFQPVYTVTDYQELIEDIVDRGRLYASENHQELDKKLKALFDV	960	Qy	601 FHQLSGMNLDAKKPPTSERQAEESGGVQGCHAENQCVSPAGTYDGAERCILCUPNG	660
Db	901 IQFKSNEBNNSARGFQPVYTVTDYQELIEDIVDRGRLYASENHQELDKKLKALFDV	960	Db	599 FHQLSGMNLDAKKPPTSERQAEESGGVQGCHAENQCVSPAGTYDGAERCILCUPNG	658
Qy	961 LAHPQNYFKYTAQESREMPRSFIRLRSKYSRFLRPyK	999	Qy	661 TFONEEGDMTCBPCPRONSGALKTPPEAWNSECGGLCQPBEYSADGFAPCOLCA-GTFQ	720
Db	961 LAHPQNYFKYTAQESREMPRSFIRLRSKYSRFLRPyK	999	Db	659 TFONEEGDMTCBPCPRONSGALKTPPEAWNSECGGLCQPBEYSADGFAPCOLCA-GTFQ	718
Qy	RESULT 3 US-1-176-847-50		Qy	721 PEGRTSFPCGGGLAKTHCGATSTFODCETRVQCSPHNTTHRCLRCVGTYOPFG	780
Sequence 50, Application US/10176847					
Publication No. US20030069636A1					
GENERAL INFORMATION:					
APPLICANT: Veiby, Peter Ole					
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR					
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST					
TITLE OF INVENTION: AND OVARIAN CANCER					
FILE REFERENCE: MRI_039					
CURRENT APPLICATION NUMBER: US/10/176-847					
NUMBER OF SEQ ID NOS: 112					
SOFTWARE: FastSEQ for Windows Version 4.0					
SEQ ID NO 50					
LENGTH: 997					
TYPE: PRT					
ORGANISM: Homo sapiens					
US-1C-176-847-50					
Query Match Score 99.6%; DB 15; Length 997;					
Best Local Similarity 99.7%; Pred. No. 0; Gaps 1;					
Matches 996; Conservative 0; Mismatches 1; Indels 2;					
RESUME 4 US-09-855-824-5					
Sequence 5, Application US/09655824					
/ APPLICANT: YAN, Chenhua et al.					
/ Publication No. US20030-66468A1					
/ GENERAL INFORMATION:					
/ TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND					
/ TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND					
/ TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND					
/ TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND					
/ FILE REFERENCE: CLO01233					
/ CURRENT FILING DATE: 2001-05-16					
/ NUMBER OF SEQ ID NOS: 6					
/ SOFTWARE: FastSEQ for Windows Version 4.0					
/ SEQ ID NO 5					
/ LENGTH: 964					
/ TYPE: PRT					
/ ORGANISM: Human					
/ TS-09-855-824-5					
Query Match: 96.8%; Score 5428; DB 12; Length 964;					

Best Local Similarity 100.0% ; Pred. No. 0; Mismatches 0;		Indels 0; Gaps 0;		Score 5056; DB 9; Length 997;	
Matches 96%; Conservative					
Qy 36 RGRAGPQEDVDEAQGLDDCHADALCQNTPTSYKCSCKPGYQCEGRGCDIDECGNELN 95					
Db 1 RGRAGPQEDVDEAQGLDDCHADALCQNTPTSYKCSCKPGYQCEGRGCDIDECGNELN 60					
Qy 96 GGCYHDLCLNIPGNYCTCFDFMLAHDGHNCIDYDECLENNGGQHTCYNVGSYECCK 155					
Db 61 GGCYHDLCLNIPGNYCTCFDFMLAHDGHNCIDYDECLENNGQHTCYNVGSYECCK 220					
Qy 156 EGEFLSDNOHTCIHRSEEGLSCMNKDHCGSHECKEAPRGSAACRGPBLAKNCRDCL 215					
Db 121 EGEFLSDNOHTCIHRSEEGLSCMNKDHCGSHECKEAPRGSAACRGPBLAKNCRDCL 180					
Qy 216 TCNHNGNGCCSISDDTAGPESCHPQYKHMHTDGRSCLEREDTLEVTSNTTSVVDCK 275					
Db 181 TCNHNGNGCCSISDDTAGPESCHPQYKHMHTDGRSCLEREDTLEVTSNTTSVVDCK 240					
Qy 276 RVKRRL-LMETCAVNNGCDCRTICKDFTSTGYHSCPYGFTIOLDGKICKD-DECQFRNGGD 315					
Db 241 RVKRRLMETCAVNNGCDCRTICKDFTSTGYHSCPYGFTIOLDGKICKD-DECQFRNGGD 300					
Qy 336 HFCKNTVGSPFCGK&GFKLTDKEKSCDQVZECSDLRTDHSCINHPGTFACACNRGVTL 395					
Db 301 HFCKNTVGSPFCGK&GFKLTDKEKSCDQVZECSDLRTDHSCINHPGTFACACNRGVTL 360					
Qy 396 YGFTHGCDTNECSINNGCQCVQCVNTGVSYECOCHPGYKLHWNKHDCKEVKGLLPTSYSP 455					
Db 361 YGFTHGCDTNECSINNGCQCVQCVNTGVSYECOCHPGYKLHWNKHDCKEVKGLLPTSYSP 420					
Qy 456 RVSLLHGKSGGGDGFPLRCHSGHLLSSDVDTIRTSVTFKLNCGSLXNALPEDEGRFA 515					
Db 421 RVSLLHGKSGGGDGFPLRCHSGHLLSSDVDTIRTSVTFKLNCGSLXNALPEDEGRFA 480					
Qy 516 LPEKHSSVKESFRYVNLTCSSGKQVPGAPRSPSKEMFTVEFELETNQKEYTASCDS 575					
Db 481 LPEKHSSVKESFRYVNLTCSSGKQVPGAPRSPSKEMFTVEFELETNQKEYTASCDS 540					
Qy 576 CIVKRETEKLRKAIRTLRKAIRHREQFHQLQSGMNDVAKKPPRTSERQAEQCGWQGHAE 635					
Db 541 CIVKRETEKLRKAIRTLRKAIRHREQFHQLQSGMNDVAKKPPRTSERQAEQCGWQGHAE 600					
Qy 636 NQCVCSSRAGTYDGAERCLCPNGTQNEQGMTCCEPGRPGNSGALKTPBEANMSEG 695					
Db 601 NQCVCSSRAGTYDGAERCLCPNGTQNEQGMTCCEPGRPGNSGALKTPBEANMSEG 660					
Qy 696 GLCQPGEYSAQDFAPCOLCALGTFOEAAGRTSCEPGGGGLATKHOGATSFQDCETRVOCS 755					
Db 661 GLCQPGEYSAQDFAPCOLCALGTFOEAAGRSCFPGGGLATKHOGATSFQDCETRVOCS 720					
Qy 756 PGHFXMTTHCIRCPVGYCPEFGKNCYSCPGNTTCDGSTMNITCCNMRGGELCD 815					
Db 721 PGHFXMTTHCIRCPVGYCPEFGKNCYSCPGNTTCDGSTMNITCCNMRGGELCD 780					
Qy 816 PTGYISPNYSGNYPANTECWTINPPKRYLILVPEIPLPIEDDCGYLWMEKTSSN 875					
Db 781 PTGYISPNYSGNYPANTECWTINPPKRYLILVPEIPLPIEDDCGYLWMEKTSSN 840					
Qy 876 SVTTYETCQYTERPIAFTSRSKLWIQFSNEGSNARGFOPYTYDQYQELIEDIVRD 935					
Db 841 SVTTYETCQYTERPIAFTSRSKLWIQFSNEGSNARGFOPYTYDQYQELIEDIVRD 900					
Qy 936 GRLYASENHOEILDKKLKIAKFDVLAHPONYFKYTQAESREMFRSFIRLRSKVSREL 995					
Db 901 GRLYASENHOEILDKKLKIAKFDVLAHPONYFKYTQAESREMFRSFIRLRSKVSREL 960					
Qy 996 RPK 999					
Db 961 RPK 964					

721 PEAGRTSCFPCCGGLATHQGATSFQDCTETRIVCSPGHPYNTTHCIRCPVGYYQPEFG 780
 :
 Db 729 PDGRTSCLSGGGLPTHLGAITSFQDCTETRIVCSPRHFYNTTICRGTCRGLGYQPEFG 776

781 KNNCVSCPNTTDFDGTSTNITOCKNRGCGELGOFTGYIESPNYPGNYPANTECTWIN 840
 :
 Db 779 KNNCVSCPNTTDFDGTSTNITOCKNRGCGELGOFTGYIESPNYPGNYPANSECTWIN 839

841 PPPKRILIVPFLPIEDDCDYLMRKTSNSNTTYCOTYERPAFTSESKKLW 900
 :
 Db 839 PPPKRILIVPFLPIEDDCDYLMRKTSNSNTTYCOTYERPAFTSESKKLW 898

901 IQFKSNEGNSARGFQVPTVYDQELIEDIVRDGLYASNEQE:LADMKL1KALFDY 950
 :
 Db 899 IQFKSNEGNSARGFQVPTVYDQELIEDIVRDGLYASNEQE:LADMKL1KALFDY 958

961 LASHQNYFRTYQESREXPFSRFLRKSRSRFLRKY 999
 :
 Db 959 LASHQNYFRTYQESREXPFSRFLRKSRSRFLRKY 997

RESULT 6

US-09-855-824-6
 ; Sequence 6, Application US/09855824
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 ; TITLE OF INVENTION:
 ; FILE REFERENCE: CLO01233
 ; CURRENT APPLICATION NUMBER: US/09/855,824
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 957
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-09-855-824-6

Query Match 88.0%; Score 4935; DB 12; Length 957;
 Best Local Similarity 90.2%; Prod. No. 0;
 Matches 862; Conservative 46; Mismatches 48; Indels 0; Gaps 0;

Qy 44 EDVDECAQCLDDCHADALCINTPSKCSKGKQYQGEGRQEDIDEKGNGCYHDCL 103
 Db 2 EDVDECAQCLDDCHADALCINTPSKCSKGKQYQGEGRQEDIDEKGNGCYHDCL 6:
 Qy 104 NIPGNYRCTCFDGFMLAHDGHNNCLDVECLENNGCQHCOHTCNGYKEGFFLSEN 163
 Db 62 NIPGNYRCTCFDGFMLAHDGHNNCLDVECLANGGCHICHTNIGSYECKEGF_FSDN 221
 Qy 164 QHTC1HRSEGLSMCNKDHGCSHICKEARPGSVACECRPFELAKNQDCTLCNHNGG 223
 Db 122 QHTC1HRSEGLSMCNKDHGCSHICKEARPGSVACECRPFELAKNQDCTLCNHNGG 181

Qy 224 COHSDCTADGPESCHPOXKHTDGPSCLERDTVELETSNTSIVDGEKRYKBRLLM 283
 Db 182 COHSCTADGPESCHPOXKHTDGPSCLERDTVELETSNTSIVDGEKRYKBRLLM 241

Qy 284 ETCAVNGGCDRTCKDTGTHSCPVGFTLQDGKTKOIDECDTRNGGDHFECNIVG 343
 Db 242 ETCAVNGGCDRTCKDTGTHSCPVGFTLQDGKTKOIDECDTRNGGDHFECNIVG 301

Qy 344 SFDCGCKKGFKLTDKECSQDVBCLSDRTGSCPIGFTLQDGKTKOIDECDTRNGGDHFECNIVG 403
 Db 302 SFDCGCKKGFKLTDKECSQDVBCLSDRTGSCPIGFTLQDGKTKOIDECDTRNGGDHFECNIVG 361

Qy 404 TNECISNNNGGQQVYCNNTGYSYEQCHGKYLHKWKKOCVVKGLPLTSVSPRVLHCK 463
 Db 362 TNECISNNNGGQQVYCNNTGYSYEQCHGKYLHKWKKOCVVKGLPLTSVSPRVLHCK 421

464 SG23DGFPURCHSGTHLSDSVTITTSVTFXLFNGKCSXNAELFPEGSRALPEKHSSY 52;
 Db 422 SG23DGFPURCHSGTHLSDSVTITTSVTFXLFNGKCSXNAELFPEGSRALPEKHSSY 481

Qy 524 KESFRYUNLTCSGKQVPGAPGRPSTPKENFITVBFELJETNCKEVTAASCOLSCIVKRTK 583
 Db 482 KESFRYANLTCSGKQVPGAPGRPSTPKENFITVBFELJETNCKEVTAASCOLSCIVKRTK 54:
 Qy 584 RLRKAINTLRLKAVRHRQFLQSGNLDAKKPPPTSEQAESCGQHAEHQCVSRA 643
 Db 542 RLRLKALPZMLZLKAHRQFLQSGNLDAKKPPPTSEQAESCGQHAEHQCVSRA 601

Qy 644 GTYDGZGWERC:LCPNGTQNEEGQNTCEPDRPNSGKJTPAWNMSBGGLUQPGCY 733
 Db 632 GTYDGZGWERC:LCPNGTQNEEGQVTCEPDRPNSGKJTPAWNMSBGGLUQPGCY 66:
 Qy 704 SADGFAPQQLCALGTQPEAQT:SCPCGGGLATHRGATSQDGETRYQCSFGHFYNT 763
 Db 662 SADGFAPQQLCALGTQPEAQT:SCPCGGGLATHRGATSQDGETRYQCSFGHFYNT 721

Qy 764 THRCIREPVGTYQPERGKNNCVSPRNNTIDFGDSTNITICKNKRGCGLGDFGTYIESP 823
 Db 722 THRCIREPVGTYQPERGKNNCVSPRNNTIDFGDSTNITICKNKRGCGLGDFGTYIESP 781

Qy 824 NYFGNYNPANTCTKTINPPKRRLIVPPELFLIEPDGCDYLMRKTSSSNSNTTYEC 883
 Db 782 NYFGNYNPANSETWTINPPKRRLIVPPELFLIEPDGCDYLMRKTSSSNSNTTYEC 841

Qy 884 QYERPLAFTSSESKKLW:QFKSNEGNSARGFQVPTVYDQELIEDIVRDGLYASEN 943
 Db 842 QYERPLAFTSSESKKLW:QFKSNEGNSARGFQVPTVYDQELIEDIVRDGLYASEN 901

Qy 944 HQEIJDKKL1KALFDYVA:HQPNYFRTYQKTAQESREXPFSRFLRKY 999
 Db 902 HQEIJDKKL1KALFDYVA:HQPNYFRTYQKTAQESREXPFSRFLRKY 957

RESULT 7

US-09-855-824-2

Sequence 2, Application US/09855824
 Publication No. US20030166048A1
 GENERAL INFORMATION:
 APPLICANT: YAN, Chunhua et al.
 TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
 NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 TITLE OF INVENTION: JSES THEREFC
 FILE REFERENCE: CLO01233
 CURRENT APPLICATION NUMBER: US/09/855,824
 CURRENT FILING DATE: 2001-05-16
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 988
 TYPE: PRT
 ORGANISM: Human
 US-09-855-824-2

Query Match 65.6%; Score 3675.5; DB 12; Length 988;
 Best Local Similarity 62.4%; Prod. No. 1-4e-27C; Gaps 14;
 Matches 654; Conservative 113; Mismatches 162; Indels 119; Gaps 14;

Gy 10 GAA---WAVULLLJLPLPLLLLAGVPPGRRAAG----POEVYDSCAQGLDQDCHADL 61
 Db 2 GALVWRWHLCVLLA-----GTGRLAQSGLPGSVVDCECSGSTDCH13: 49

Qy 62 CONTPSYKCSCKKGFKLTDKECSQDVBCLSDRTGSCPIGFTLQDGKTKOIDECDTRNGGDHFECNIVG 120
 Db 50 CONTPSYKCSCKKGFKLTDKECSQDVBCLSDRTGSCPIGFTLQDGKTKOIDECDTRNGGDHFECNIVG 109

Qy 221 HQCHNLCDYDCELENNNGCQHCTCNCVMSYEQCHGKYLHKWKKOCVVKGLPLTSVSPRVLHCK 180
 Db 110 HQCHNLCDYDCEQNNNGCQHQVYCNNTGYSYEQCHGKYLHKWKKOCVVKGLPLTSVSPRVLHCK 69

Qy	852	PEIFLPIEDDCGDYLYMVRKTSSNSVTTYETCOTYERPIAFTSSKKLWICFSNEGNSA	91;
Dtb	841	PEIFLPIEDDCGDYLYMVRKTSSNSVTTYETCOTYERPIAFTSSKKLWICFSNEGNSG	900
Qy	912	RGFQVPPVYTYDQXQELIEDIVRDGRYLAYSENHOIUKKLKALFDVLAKHONYFXYT	971
Dtb	901	KGFFQVPPVYTYDQXQELIEDIVRDGRYLAYSENHOEIKDKKLKALFDVLAPNHFXYT	960
Qy	972	AQEREMPPRSFRLLPSKVSRLPBYK	999
Dtb	961	AQESEMPPSFYKLLPSKVSRLPBYK	988
RESULT 9			
	US-09-898-570-14	Sequence 14, Application US/09898570	
	Patent No. US2002013612A1	Patent No. US2002013612A1	
	GENERAL INFORMATION:		
	APPLICANT: GERLACH, VALERIE L.		
	APPLICANT: EULERYAN, KAREN		
	APPLICANT: MACDONALD, JOHN R.		
	APPLICANT: SMITHSON, GLENDA		
	TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USING THE SAME		
	FILE REFERENCE: 15966-776CIP		
	CURRENT APPLICATION NUMBER: US/09/898, 570		
	CURRENT FILING DATE: 2001-07-03		
	PRIOR APPLICATION NUMBER: 60/198, 293		
	PRIOR FILING DATE: 2000-04-19		
	PRIOR APPLICATION NUMBER: 60/198, 645		
	PRIOR FILING DATE: 2000-04-20		
	PRIOR APPLICATION NUMBER: 60/220, 809		
	PRIOR FILING DATE: 2000-06-09		
	PRIOR APPLICATION NUMBER: 60/199, 476		
	PRIOR FILING DATE: 2000-04-26		
	PRIOR APPLICATION NUMBER: 60/200, 025		
	PRIOR FILING DATE: 2000-04-26		
	PRIOR APPLICATION NUMBER: 60/224, 610		
	PRIOR FILING DATE: 2000-08-12		
	PRIOR APPLICATION NUMBER: 60/200, 024		
	PRIOR FILING DATE: 2000-04-26		
	PRIOR APPLICATION NUMBER: 60/199, 080		
	PRIOR FILING DATE: 2000-04-26		
	PRIOR APPLICATION NUMBER: 60/218, 591		
	PRIOR FILING DATE: 2000-07-17		
	PRIOR APPLICATION NUMBER: 60/271, 814		
	PRIOR FILING DATE: 2001-02-27		
	PRIOR APPLICATION NUMBER: 60/215, 055		
	PRIOR FILING DATE: 2000-07-03		
	PRIOR APPLICATION NUMBER: 09/839, 446		
	PRIOR FILING DATE: 2001-04-19		
	NUMBER OF SEQ ID NOS: 58		
	SOFTWARE: PatentIn Ver. 2.1		
	SEQ ID NO: 14		
	LENGTH: 974		
	TYPE: PRT		
	ORGANISM: Unknown Organism		
	FEATURE:		
	OTHER INFORMATION: Description of Unknown Organism: PO_YX		
	OTHER INFORMATION: 297832_B_1		
	US-09-898-570-14		
Qy	22	LPPILLAGAVPPGRGAAQPQEDVDEAQGLDDCHADLCTNTSYKCSCKPQYQGBG	81
Dtb	6	VPGCLLVLVHARAQSKAAQDVCEVGTNDCHDAICNTPSYKJCKSGSYT3G	65
Qy	92	RQCED1DCEGNELNGCYHDCLN1PGNYRCTCFDGFMLAHGHNCDLVDCELNNGCCG	141

Db	RHKDVKDECEREDNAGSVHDCVN:PNYRCTCYGPHLAHGHNCLDVEAEGNGCQQ	125
Qy	TCVNMGSYECCKESPFELSDNCHTICHRSEBGSMMKDHGCSHICKEAPRGSVACECR	261
Ds	SCVNMGSYECHECREFFESNCHTICRPEEGMNKNKGCAHCIRETPGGTAECR	195
Qy	PGFELAKNQRDCLTGNHNGCQHSACDADPESCHPQYKHTDGRSCLERETVLE	261
Ds	PGFELKNCADCKLTONYGNNGCQHCDDEQGPREGCHIKFVLPDGKTS-----	237
Qy	VTESNITSVWDPKYKRR-----LNEYCAVNNGCQRTKSTSVHCSCPVGFTLQL	326
Ds	-----GERRELCIPTQAVSNETCAVNNGCDSKHDAAQVHCTCPVGMPLQP	286
Qy	DGKTKDIDEOTRNNGCDDHFCNTNGSPDGCCKGSKFLLTDEKSQDVDECSJCTCDH	376
Ds	DRKTKDDECRLNNNGCDDHICRNTGFSFECSCKGKLLINERKVQDIECSFRICDH	346
Qy	SCINHGTFAACACNRGTYLGYFTHCGDTNECSINNGCCQCVNTGSYEFOCHPQY-KL	435
Ds	SCYNTTGFSCQCLCHRATYLLGYTHCGDVDECSINGRCFGFQSYCQCPAGSGRL	406
Qy	HWNKCKCVCYEVKOLPPTSVSPRVLSSCKSGGDGCCFLRCHSGHL--SSDYTTIRTSVF	493
Ds	HWNKGKOCTEPLKCGGSAGASMLSCNRSGKDKTCAALTCPSRARFLPEAAVLSIKQASF	466
Qy	KLINEGKC-.SLXNAELPPEGURAJPJXHSSVYKESTRYVNLTCSKQVCPGAPGRSTPK	551
Ds	6467 XIKDAKRLHRNKGKTEEAAGRITGGAGAPSECQVTFILKCDSSRKGKERRARTPPGK	526
Qy	EMF--TTEFELETNOKEVATASDLSCIVKRETRKAHREQFHPLCLSRNL	610
Ds	527 EVTRTIELEAVRAEETTASGCP-PRQERRRLIGSLACMRKSNQCREFLRLAGLDY	586
Qy	DVAKKPPRTSERQA--ESCGVGOSHBBENOCVSCRAGTYDGAERECILCPNGTFCNEEG	667
Ds	587 EAHKPLVAGRAEYPMESCRGQHAGTCKVSCP2OYTYYISQ-TECVPFCAGTFFOREG	646
Qy	QMTCEPPRPNGSAKLTKPTPEAMMSGGGLCPGEYSADGFPQCGCALGFOPEAGRTS	727
Ds	647 QLSCDUCPSDADGPL---GATNVTAGCOPGSSVGDGFKPCOPSPRGTYQPEACRTL	703
Qy	CPFGGSGATKQGATFQCGTTRVQSPGHYNTTHRCRCPVGYTOPEGKKNVQVSC	787
Ds	704 CPFGGSGATKQGATFQCGTTRVQSPGHYNTTHRCRCPVGYTOPEGKKNVQVSCSRC	763
Qy	PGNNTBDGGSNTITQCNRGGELGDFGTGIESPYPNYGNPANTECTWT-NPPKBRZ	847
Ds	764 PGNNTBDGGSNTITQCNRGGELGDFGTGIESPYPNYGNPANTECTWT-NPPKBRZ	823
Qy	LIVYPEIFLPIEDCGY-YMKRTSSNSVITYETCOTYERPLAFTSRSKKWLWQFSNE	907
Ds	824 LIVYPEIFLPIEDCGY-YMKRTSSNSVITYETCOTYERPLAFTSRSKKWLWQFSNE	893
Qy	GNSARGPQVYTYDQWQELLEUDGRYLASENQEI...KDKKLKALPEWLMHQNY	967
Ds	884 ANSARGPQVYTYDQWQELLEUDGRYLASENQEI...KDKKLKALPEWLMHQNY	943
Qy	FKYTAQBSREMPERFIRLJRSKVSFLRPYK	999
Ds	944 FKYT-ERKEMDPKSFYKLRSKVSFLRPYK	974

; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/639,446
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 14
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; US-09-639-446-14

Query Match 61.0%; Score 3417.5; DB 11; Length 974;
Best local Similarity 60.1%; Pred No. 5,7e-25; DB 11; Length 974;
Matches 596; Conservative 140; Mismatches 219; Indels 37; Gaps 9;

QY 22 LPPPLLLAGAVPPGRRAAKPQEDCQGLDCHADALCONTPSYKSCPKPGCAG 81
Db 6 VPSCLLVLVHARAQYSKRAQDDECVEGTGTCRHDICATPPSYKICKSGTGTG 65
QY 92 ROCEDIDBEGNBELNCSVCDLNITPFGNRTCTEDGFMALAHCHNCUDSCLENNGGCG 141
Db 66 FKCKDVBCEEREDNAQCVHDVNIFGNRCTCYDGFLAHGNCJSDCAEGNGGCQ 125
QY 142 TCVNVGSKYSECKGSPFSDNOITCIRSEEGLSCMKDKGSHCKEPRGSACER 201
Db 126 SCVNMMSYECHECRGEFFJSDNQBTICPEEGNMCMNKHCQAHCRESFKGACER 185
QY 202 PPFELAXNQRQCILCTNHNGGCGHSDDTADPECCSCHPCKMHDGRSCLEBDTUE 261
Db 186 PSFELTQCRDCKLTKYNGGCGHTCDDTECGPRCGCHIKFVHLIDGKCI----- 237
QY 262 V-EESTNTSVWDGDKRKR-----IMECAVNGGCDRTCKTGVHRSCKPFGQI 316
Db 238 -----OBRJECQHTPOAQWNETCAVNGGCGDOSKCRDACTGVCHTCPVGFNLQP 286
QY 317 DCKTKDIDEQTRGSCDFCKRQVUGSDFCGKKGKLUDEKSCCOVCECSRJTC 376
Db 287 DKTKDIDEQTRGSCDFCKRQVUGSDFCGKKGKLUDEKSCCOVCECSRJTC 346
QY 377 SCNHPGTFACAGNGYTGFTGDPGNECSNNGGCOVNTVSSVQCHQGYKL 435
Db 347 ICPNTPOSGQCLCRRYLIGITGQDVBCSINRGGRCGINTPSYCTCAGQR 406
QY 436 HNKDKDVEVGKLLPSVSFRVSHGSGGGCFACQH--SSUTTIRSGRI 433
Db 407 HNKDKDTEPLKQCGSPGASKAMLSNRGSKOTCALTOPSPARFPEAVLISKORASF 466
QY 494 KUNGKC---SIKNAELPPEGIRPALPEKHSSVKEFRVNUTCOSKQVGAAPRSPK 551

Query Match: 60.6%; Score 3397; DB 12; Length: 993;
Best local Similarity 56.3%; Pred No. 2,1e-249; DB 12; Length: 993;
Matches 595; Conservative 136; Mismatches 212; Indels 73; Gaps 9;

QY 22 LPPPLLLAGAVPPGRRAAKPQEDCQGLDCHADALCONTPSYKSCPKPGCAG 81
Db 6 VPSCLLVLVHARAQYSKRAQDDECVEGTGTCRHDICATPPSYKICKSGTGTG 65

Patent No. US2002012362A1
 GENERAL INFORMATION:
 APPLICANT: GERJACH, VALERIE L.
 APPLICANT: ELJERMAN, KAREN
 APPLICANT: MACDOUGALL, JOHN R.
 APPLICANT: SMITHSON, GLENDA
 TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USING THE SAME
 Little of Invention: METHODS OF USING THE SAME
 FILE REFERENCE: 15966-776CIP
 CURRENT APPLICATION NUMBER: US 09/898,570
 CURRENT FILING DATE: 2001-07-03
 PRIOR APPLICATION NUMBER: 60/198,293
 PRIOR FILING DATE: 2000-04-19
 PRIOR APPLICATION NUMBER: 60/198,645
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: 60/210,809
 PRIOR FILING DATE: 2000-06-09
 PRIOR APPLICATION NUMBER: 60/199,476
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/200,625
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/224,610
 PRIOR FILING DATE: 2000-08-11
 PRIOR APPLICATION NUMBER: 60/200,324
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/199,800
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/218,591
 PRIOR FILING DATE: 2000-07-17
 PRIOR APPLICATION NUMBER: 60/271,914
 PRIOR FILING DATE: 2001-02-27
 PRIOR APPLICATION NUMBER: 60/215,555
 PRIOR FILING DATE: 2003-07-03
 PRIOR APPLICATION NUMBER: 60/219,446
 PRIOR FILING DATE: 2001-04-19
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 16
 LENGTH: 10C9
 TYPE: PRT
 ORGANISM: Unknown Organism
 FEATURE: OTHER INFORMATION: Description of Unknown Organism: POLYX
 OTHER INFORMATION: CGG5096-04
 US-59-898-570-16

Query Match Best Loca_ Similarity 60.6%; Score 3394; DB 10; Length 1039;
 Matches 597; Conservative 142; Mismatches 216; Indels 72; Gaps 20;

750 HFNNTTHRC-RCPVTYCPCFGKONCVSPGNNTDFGSNTNICKNRCCGGEGFT 817
 753 FVNTSIHRCAMGSYQDFRQNFCSRCPGNTSDPQGSTSWAQCNQRQCGEIGFT 812

818 GVEESPVPGWPNPANTECTWTPPPKKRLIVWEIPLIEDDGDVWVKRSSNSV 877
 813 GVEESPVPGWPNPANTECTWTPPPKKRLIVWEIPLIEDDGDVWVKRSSNSV 872

878 TTYETCYTERPIAFFRSKLUWIFSKNEGSNSARGFOVPUVTDYGLIEZIVRGR 937
 873 TTYETCYTERPIAFFRSKLUWIFSKNEGSNSARGFOVPUVTDYGLIEZIVRGR 932

938 LYASENHOEILKKLUKALFDVLAHQNPYKTAQESREMPSFIRLRSKYSRFLRP 997
 933 LYASENHOEILKKLUKALFDVLAHQNPYKTAQESREMPSFIRLRSKYSRFLRP 992

998 YK 999
 Db 992 YK 993

RESULT 12
 US-09-898-570-16
 Sequence 16, Application US/09898570

Qy 377 SCINHPTFACACNRGTYLGFTHGDZNECSZNGGCQCVNTVSYECOHPGY-KL 435
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Db 347 ICVNTPGSFOCLCHRGYLYGITHCGDVBCS.NRGCRFGCINTPSYQCTCPAGQRL 406
 Qy 436 HNKKDCVEVKGLPLTSVSPRSVHCKGGGGGCGFLRC----- 474
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Db 407 HNGKDCTEPKCQSPGASKAMLSNRGSKKRTCALCPSRFLPESENGFTVSGTP 466
 Qy 475 -----HSGHLLSSD-----VTIRTSVFKLNGEKC-SLKNAELPEEG-RPAL 516
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Db 467 SPRAPARAGHNGNTRSNHCAAVSLIKRASPKIKDAKCRHLRKNGKTEAGRITG 526
 Qy 517 PERKSSVKEFSRYYNLTCSGSKQPGAPRPSKEMF-IVFELETNQEVTCASCLS 575
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Db 527 PGGAPSECQVTFIHKCDSRKGRRARTPPKEVTRJTLAEVRAETTASCLP 586
 Qy 576 CIVRMRERLKGSJMLRKSINCJFJLALGJYEAHKPGVAGERAEPMESCRQH 646
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Qy 633 HAENQCVSFRAGTYTDGARERCICPNGFONEBGMQTCPCPRPGNSGALKP-EPAWMS 692
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Db 647 RAGTKCVSPGPSTYHGQTEQCVCPAGFQEREGQLSDLPSSDANGPL--GATNVT 703
 :||:||:||:||:||:||:||:||:||:||:||:
 Qy 693 ECGGICQOPBYSAGFAFCOLCAAGTQFCEAGRSFCGGGATKGATGATPSQDCETRY 752
 :||:||:||:||:||:||:||:||:||:||:||:
 Db 704 TCAQOCPPGSHVSGFKPCQCPRTYQPEAGRITLCFPGGGL-KHEGAISPOCDOTV 763
 :||:||:||:||:||:||:||:||:||:||:
 Qy 753 QCSQHYYNFSIHCIRCAMSQYDPERQNCSCPQHNSDJDGSTSVACQCNRQCGE 823
 :||:||:||:||:||:||:||:||:||:||:
 Db 764 QCSQHYYNFSIHCIRCAMSQYDPERQNCSCPQHNSDJDGSTSVACQCNRQCGE 823
 :||:||:||:||:||:||:||:||:||:
 Qy 813 LGDFDTGYLISPNNYQNYPANTECTWTINPPKRLIVPEIRPLIEDCGDVMRKT 872
 :||:||:||:||:||:||:||:||:||:
 Db 824 LGEGTGYLISPNNYQNYPACVETANINPPKRLIVPEIRPLSEDCCGJUNRNS 983
 :||:||:||:||:||:||:||:||:
 Qy 873 SSNVTTYEROTPIERTSRKKLUQFKSEGNARGFQPYVYDQEQLIED 932
 :||:||:||:||:||:||:||:||:
 Db 884 SPSSITYEROTPIERTSRKKLUQFKSEGNARGFQPYVYDQEQLIED 943
 :||:||:||:||:||:||:||:
 Qy 933 VDRDLYASMEHOLKDKLXIAFLVLAHQWQYKTAQSEREMERSERFLRSKV 992
 :||:||:||:||:||:||:||:
 Db 944 VDRDLYASMEHOLKDKLXIAFLVLAHQWQYKTAQSEREMERSERFLRSKV 1002
 :||:||:||:||:||:
 Qy 993 RFLRPIK 999
 :||:||:
 Db 1003 SFIREYK 1009
 :||:
 RESULT 13
 JS-09-839-446-16
 ; Sequence 16, Application US/09839446
 ; Publication No. US20030050232A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GERACH, VALERIE L.
 ; APPLICANT: ELLERMAN, KAREN
 ; APPLICANT: MCDougall, John R.
 ; APPLICANT: SMITHSON, GLENDA
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
 ; METHODS OF USING THE SAME
 ; FILE REFERENCE: 15966-776
 ; CURRENT APPLICATION NUMBER: US/09/839,446
 ; CURRENT FILING DATE: 2003-04-19
 ; PRIOR APPLICATION NUMBER: 2000-05-29
 ; PRIOR FILING DATE: 2003-04-19
 ; PRIOR APPLICATION NUMBER: 60/198,645
 ; PRIOR FILING DATE: 2003-04-20
 ; PRIOR APPLICATION NUMBER: 60/210,809
 ; PRIOR FILING DATE: 2000-05-09
 ; PRIOR APPLICATION NUMBER: 60/199,476
 ; PRIOR FILING DATE: 2000-04-26
 ; PRIOR APPLICATION NUMBER: 60/200,025
 ; PRIOR APPLICATION NUMBER: 60/200,025
 PRIOR FILING DATE: 2000-04-26
 PRICR APPLICATION NUMBER: 63/224,610
 PRIOR FILING DATE: 2003-05-11
 PRIOR APPLICATION NUMBER: 63/200,024
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 60/199,880
 PRIOR FILING DATE: 2000-04-26
 PRICR APPLICATION NUMBER: 60/218,591
 PRIOR FILING DATE: 2000-07-17
 PRIOR APPLICATION NUMBER: 60/271,814
 PRIOR FILING DATE: 2001-02-27
 NUMBER OF SEQ ID NOS: 49
 SEQ ID NO: 16
 LENGTH: 1009
 TYPE: PTM
 OTHER INFORMATION: Description of Unknown Organism
 ORGANISM: Unknown Organism
 FEAT: JRE
 OTHER INFORMATION: Description of Unknown Organism: POLYX
 JS-09-839-446-16
 Query Match Score: 3394; DB: 1; Length: 1009;
 Best Local Similarity: 58.1%; Pred No.: 3.6e-249;
 Matches: 597; Conservative: 142; Wmismatches: 216; Indels: 72; Gaps: 10;
 Qy 22 IAPILLJAGAVPPGRAGQPQEDBECKQDGDCHADLCQNTTSYKSCSKCYQGEG 81
 :||:||:||:||:||:||:||:||:
 Db 6 VPGCLLVLVHLHARAQSKAQDVPDCECTEGTDNCHIDACONTPRSYKICKSYGTGQ 65
 :||:||:||:||:||:||:||:
 Qy 82 RQCEQIDECGHNGLNGCVRHPCINTPNYRCTCFCDPFLWALDGHNNTLDVZBCCJENNGCQH 141
 :||:||:||:||:||:||:||:||:
 Db 66 HNKDCVDECEBREDNAGCVCBDCVNPNQYRCTCYDGFLHNDGHNLDVBEAEGNGGQO 125
 :||:||:||:||:||:||:||:
 Qy 142 TCVNNGSYBCCKGSPFLDNHCHISEGGSQMNKDGCSICKAPRGVACECR 201
 :||:||:||:||:||:||:||:
 Db 126 SSVNMGSYEHCREGFFLSDNQHCTIOPEEGMCMNKHGCAHICREFFGG-AECCR 165
 :||:||:||:||:||:||:||:
 Qy 202 PGEFLAKNQRCILQHNGNGCQHSCTDAGPSCSCHQKOMTGPSCLEREDTV 261
 :||:||:||:||:||:||:||:
 Db 186 PGFELTNORBCKLCTNYGNGQGQFCDDPSEGPRGCCHIKFVLIHTDGKTCI----- 237
 :||:||:||:||:||:||:||:
 Qy 262 VTESNTISWVGDKRVRRL-----LMETAVANNGCDRTCKDNTGVHRCSCPVSFTQL 316
 :||:||:||:||:||:||:||:
 Db 238 -----GERILLEQHTPQAVSNETGAVANGGDSKHDQHTGVTGHTCPVFMQLP 286
 :||:||:||:||:||:
 Qy 317 DCKTKDQIDEOTRNGGDFHCKN-VGSFPGCKKGKFKLTDKSKDQDVBCSLDRTCH 376
 :||:||:||:||:||:||:||:
 Db 287 DRKCKDQIDEOTRNGGDFHCKN-VGSFPGCKKGKFKLTDKSKDQDVBCSLDRTCH 346
 :||:||:||:||:||:
 Qy 377 SCINHPTFACACNRGTYLGFTHGDZNECSZNGGCQCVNTVSYECOHPGY-KL 435
 :||:||:||:||:||:||:||:
 Db 347 ICVNTPGSFOCLCHRGYLYGITHCGDVBCS.NRGCRFGCINTPSYQCTCPAGQRL 406
 Qy 436 HNKKDCVEVKGLPLTSVSPRSVHCKGGGGCGFLRC----- 474
 :||:||:||:||:||:||:||:
 Db 407 HNGKDCTEPKCQSPGASKAMLSNRGSKKRTCALCPSRFLPESENGFTVSGTP 466
 :||:||:||:||:||:||:
 Qy 475 -----HSGHLLSSD-----VTIRTSVFKLNGEKC-SLKNAELPEEG-RPAL 516
 :||:||:||:||:||:||:
 Db 467 SPRAPARAGHNGNTRSNHCAAVSLIKRASPKIKDAKCRHLRKNGKTEAGRITG 526
 :||:||:||:||:||:
 Qy 517 PERKSSVKEFSRYYNLTCSGSKQPGAPRPSKEMF-IVFELETNQEVTCASCLS 575
 :||:||:||:||:||:
 Db 527 PGGAPSECQVTFIHKCDSRKGRRARTPPKEVTRJTLAEVRAETTASCLP 586
 :||:||:||:||:
 Qy 576 CIVRMRERLKGSJMLRKSINCJFJLALGJYEAHKPGVAGERAEPMESCRQH 632
 :||:||:||:||:
 Db 587 CIVRMRERLKGSJMLRKSINCJFJLALGJYEAHKPGVAGERAEPMESCRQH 646
 :||:||:||:
 Qy 633 HAENQCVSFRAGTYTDGARERCICPNGFONEBGMQTCPCPRPGNSGALKP-EPAWMS 692
 :||:||:||:
 Db 647 RAGTKCVSPGPSTYHGQTEQCVCPAGFQEREGQLSDCPGSDAHGPL--GATNVT 703

QY 693 ECGGCGCGEYADGFAPCQCALGTQPEAGRTSPCGGGLATHGQAGNSQQCETRV 752
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 704 TCGGCPGPGOHSVDDGKFCQCPGRGYQPEAGRTLPCGGGLTRHREGAISFOCDTK 763
 QY 753 QCGPGHVNNTTHRCIRCPGVYQPERGKANVSCONTTDFDSTNITOCRNRGGE 812
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 764 QCGPHYNTSHRCIRCAMGSYQDPFRQFCSRCPGNSTFDGTSVACCKNRCCGE 823
 QY 813 LGDFTGTTESPNVPGNPANTCTWTINPPPKRRIIVPFLPILEDGGDLWMRKS 872
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 824 LGSFTGTTESPNVPGNPAGVCIWINNPPKRLIVPFLPSEDECGLVWRKS 883
 QY 873 SSNSVTVTETCOTYPEPIAFTRSKWLIQFSNEENSARGPQVPTYDQIREDI 932
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 884 SSPTTETCOTYPEPIAFTRSKRULINPKTSEANSARGPQIPTVTDYDQIREDI 943
 QY 933 VDGRLYASENHOELKKKLIKAFLVLAHQNYFVYQGERSRENPRSTRILASKV 992
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 944 VDGRLYASENHOELKKKLIKAFFEVLAHQNYFVYQGERSRENPRSTRILASKV 1002
 QY 993 REFLPYK 999
 Db 1003 SELRPYK 1009

RESULT 14
 US-10-239-663-58
 ; Sequence 58, Application US/10239663
 ; Publication No. US2003C139572A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agarwal, Parkaj
 ; APPLICANT: Murdoch, Paul R.
 ; APPLICANT: Rizvi, Safia K.
 ; APPLICANT: Smith, Randall F.
 ; APPLICANT: Xiang, Zhaoying
 ; APPLICANT: Rabnick, Karen
 ; TITLE OF INVENTION: NOVEL COMPOUNDS
 ; FILE REFERENCE: GPS0018
 ; CURRENT APPLICATION NUMBER: US/10/239, 663
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: PCT/US01/09226
 ; PRIOR FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 66192, 158
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 66192, 668
 ; PRIOR FILING DATE: 2000-03-27
 ; PRIOR APPLICATION NUMBER: 66192, 166
 ; PRIOR FILING DATE: 2000-04-27
 ; NUMBER OF SEQ ID NOS: 66
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 58
 ; LENGTH: 964
 ; TYPE: PR
 ; ORGANISM: Homo sapiens
 ; US-10-239-663-58

Query Match 58.7%; Score 3290.5; DB 12; Length 964;
 Best Loca: Similarity 59.7%; Pred. No. 2.5e-241;
 Matches 580; Conservative 110; Mismatches 162; Index 119; Gaps 14;
 QY 10 GAA...-WAVLILLLPULLAGAVPGRGRAG...-PQBDVDSGAQGWDQCDRADAI 61
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 2 GAAAVRWHLCLVLLA...---GTRGLLAGGSLPGSVDVDECBSEGTDCHDAI 49
 QY 62 CQMPPTSKKOSCPGQGGRQEDIDECGNE_LNGGCVHDCLNIPGNMRCFDFGMLA 120
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 50 CQMPPTSKKLCRGKYGEGKQEDIDECNDVYNGVHECINTPNYRCTDFGMLA 109
 QY 121 HDGNCNLDVDECLENNGCQHTCVMNSYECCKKEGFLPSLDNOHTCHRSEBZGLSNK 180
 Db 11C HDGNCNLDVDECDQDNGGQQCICUMAMGSYECOCHSGFPFLSDNQHTCIRHRSNEGNGMK 169

RESULT 15
 JS-10-221-097-51
 ; Sequence 51, Application US/10221097
 ; Publication No. US20030114476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agarwal, Parkaj
 ; APPLICANT: Murdoch, Paul R.
 ; APPLICANT: Rizvi, Safia K.
 ; APPLICANT: Smith, Randall F.
 ; APPLICANT: Xiang, Zhaoying
 ; TITLE OF INVENTION: NOVEL COMPOUNDS
 ; FILE REFERENCE: GPS0016
 ; CURRENT APPLICATION NUMBER: US/10/221, 097
 ; CURRENT FILING DATE: 2002-09-06
 ; PRIOR APPLICATION NUMBER: PCT/US01/07143
 ; PRIOR FILING DATE: 2001-01-05

QY 321 STGHHCSCPVGFTQDGKCKCJDECOFRNGGCDHFKRNQVSFDGCKKGKLUDEK 360
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 320 TQDQIDECSFERTCDHICNSPSSQCLCHRGTLYGTHGCGTNECS:NGGQCVUN 420
 QY |||||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 361 SCQDVBESLDRTDHSCHINHPGTFACAGNRGTYLGHGCGTNECS:NGGQCVUN 420
 QY |||||:|||:|||:|||:|||:|||:|||:|||:
 Db 380 TKGTEBCVPPGRRLHNGKDCZTGKLSRAKTSRPAQJSCSKAGGYESCFLSPCAHL 439
 QY |||||:|||:|||:|||:|||:|||:|||:
 Db 475 --
 QY |||||:|||:|||:|||:
 Db 440 FVPSSENSVVISCGVPGPCKALQKNGTSSGGPGSCDAPTPIKOKARFKRDKACHL 499
 QY |||||:|||:|||:|||:
 Db 503 KNAELFPEGIRPAPEKHSVKESR-----YVNCGSGKQVPGAPERPSTKE 552
 QY |||||:|||:|||:|||:
 Db 500 R-----PHSQRAKETAPRQLPDHCHTFVZLKCSSKK--RRRGRKSPKE 544
 QY |||||:|||:|||:
 Db 553 X-FITVEPELETOKEVTAASCCLCIVKESTERKURKARTLRKAVRHOFHQLSGMNJD 611
 QY |||||:|||:|||:
 Db 545 VSHITAEFEETKCEEASOTCEADCLRKAQESQCAA-KTRKSIROOFYVQNSSTEY 604
 QY |||||:|||:|||:
 Db 612 VAKKPRPRSERQASCQGQGQHANQCSRACTYQGARERCICLCPNQTCQVZRGQTC 671
 QY |||||:|||:|||:
 Db 625 VAQRPAKA-EGQG-ACGAGSVAVLQDPSKVACGFGHFGGELQCVSMPGTYQONEGQLSC 653
 QY |||||:|||:|||:
 Db 672 EPCFPGNSGALKTPEAMWNSCCSGLCGGAEYISDGFACQLAIGTQPEAERTSCPC 733
 QY |||||:|||:|||:
 Db 664 TPCP--SSSGLG-PGARNVSECCQCSGFFSDGFRPCQAPVGTTQEPRTGFC 720
 QY |||||:|||:|||:
 Db 732 GGGLATKQGATSFQCDERTRVQCGPHEFNMTTRCICPVGTVOPERKNNVSCPGNT 791
 QY |||||:|||:|||:
 Db 721 GGGLATKHEGTTSFDCEAHKVHCPGHMYNTTACRICEPVGTVOPERKNNVSCPGNT 786
 QY |||||:|||:|||:
 Db 792 TTDPOSTK-TCCNRRGCGELGFTGTMESPNGPNPANTCTWTINPPKRLIVW 851
 QY |||||:|||:|||:
 Db 793 STDFGOSTNVTNTECKNHQCGCSELG3TYGTYTESPNVPGDQYDANAECVWHIAPPKRLIVW 840
 QY |||||:|||:|||:
 Db 852 PEIPLIEDCGDVLWMRKSASSNSVTVTETCOTYER-#AFTERSKKWIQFSNEGNSA 911
 QY |||||:|||:|||:
 Db 841 PEIPLIEDCGDVLWMRKSASSNSVTVTETCOTYER-#AFTERSKKWIQFSNEGNSA 900
 QY |||||:|||:|||:
 Db 912 REFLPYK 922
 QY |||||:|||:|||:
 Db 931 REFLPYK 921
 QY |||||:|||:|||:

```

; PRIORITY APPLICATION NUMBER: 60/787,107
; PRIORITY FILING DATE: 2000-03-06
; PRIORITY APPLICATION NUMBER: 60/236,874
; PRIORITY FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/188,916
; PRIORITY FILING DATE: 2000-03-13
; PRIORITY APPLICATION NUMBER: 60/237,846
; PRIORITY FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 51
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-221-097-51

Query Match: 58.7%; Score 3290.5; DB 12; Length: 965;
Best Local Similarity: 59.7%; Pred. No. 213-241; Mismatches: 110; Indels: 119; Gaps: 14;
Matches: 58C; Conservative: 110; N mismatches: 162; N indels: 119; N gaps: 14;

Qy 10 GAA---WAVELLILPPPLLAGAVPPGRGAG---PQDVDECAQGLDCHADL 61
Db 2 GAAAVRWHICVLL-----GTRGRAGGSGLPGSVDCESEGTDCHC 49

Qy 62 CQNTTSYKSCCKQGYQSGRQEDIDECKNE LANGCYFDCUNIPGNVRCTCPGEMLA 120
Db 50 CQNPKPSYKICKPGYKGSKQCDIDECANDYNGGCVHECINIPGNVRCTCFDFYLA 109

Qy 121 HDGNCNLDYDECILNGGQHCTCVNNMSYECCKEKGEPFLSDCCTCHRSSEGLSYK 180
Db 110 HDGNCNLDYDECQNGGQHCTCVNNMSYECCKEKGEPFLSDCCTCHRSSEGLSYK 169

Qy 181 DHGSHICKEAPRSSVACCRPGEELAKNQRDILTCMHNGGQHSCDTADGPECSCH 240
Db 170 DHGCAHICRETPKGAVACDRCRPGFLDIAGNOKDCTLTCHNGGQHSCDDTGTGCH 229

Qy 242 PQYKHMHTGRSCJCREDTVLETVSNTISWDGDKRVRRLMTCAYVNNGGDRTKDT 300
Db 230 QKVALHSDERTCL-----EITCAYNNGGCDRICKC 259

Qy 301 STGYNHSCCPVGGFTIQLDGKTKDIDECOFRNGGCDHFCRNTVGSFDGCKKPKJLDEK 360
Db 260 ATGRCSCCNVFTIQLDGKTKDIDECOFRNGGCDHFCRNTVGSFDGCKKPKJLDEK 360

Qy 361 SCQVQDECSDLRCDHSCNHPGTFACAGNRGTYLGTHCGTNEESINNGGQQCVN 420
Db 320 TCGQIDECSPERTCDHICINSPGSPQCLCHRGYLYGTHCGDWDCECSMSKSCDDCVN 379

Qy 421 TVGYPECCQCHPGYKLHNWKDCVEV-KGLPFTSVPRIHLHCGSKGGGCPIC--- 474
Db 380 TKGSYEVCPGPGRRLHNGKDCVETGKCLSRAKTSRPAQSLCAGGVSCLSPRAHTL 439

Qy 475 -----HSGHHS---SDVTITRSVTFKLNNEKCSJ 502
Db 440 FVPPSENSYVLSGVSGPGQCKALOKANGTSSGLGPGSCADPTPIKOKARFKRDKAHL 493

Qy 503 KNAELPFEGLRPAJPEKHSVKSEFR-----YNTLTCSSGKOPGAPGRSPPK 552
Db 500 R-----PHSQARAKETARQFELDHDCHVTPVILKCDSSKA--RR2GRKSPSK 544

Qy 553 M-FITVEFPLETNKEVTAASCDSLUSCIKIRTEKXKIRTAKAVHREQFHQDLSGMLD 621
Db 545 VSHTAEELETKHEASDTCEADCLRKRAEQSQAATLKRSQGROOFYVQWSTEYE 604

Qy 612 VAKPPTTSERQASCGVGQHANQCUSCRAGTYDGAERCLCICNGTFRNEBGMTC 671
Db 603 VAQRPKALEGQG-ACGAGQVLODSKCVACGSPSTHFGS;GQGVSCHMPGTYQDMEQJSC 663

Qy 672 EPCCRPGNCSALKTPEAWNMSECGLCOPGEYVADGFAPCQLCALGTFOPENGRTSFPC 732
Db 664 *TPCP--SSDGJGPGARNVSECGQCSGFFSADGFRPCQACPVGTYQPERGTRGCP 720

Qy 732 GGGLATKHGATSTSFQDCE-RVQCSGPHFYNTTTRCIRCPVGVYQPRBRGKNCVSCPNT 791

```

```

Search completed: October 22, 2003, 11:40:28
Job time : 97 secs

Db 721 GGGLATKHGATSTSFQDCEAKVHSPGHNTTHRCCRCPVGTQPERGQHNCITCPNT: 780
Qy 792 *TDPGSTNTQCINRRCGELDGTGYESPYVPGNUPANPCTWTNPPCRILW 851
Db 781 STDPGSTNTQCINRRCGELDGTGYESPYVPGNUPANPCTWTNPPCRILW 840
Qy 852 PEIPIPIEDGDMWRYKTSSSNSSVTVETCOTYERIAFSSRSKWIQFNSNEGNSA 911
Db 841 PEIPIPIEDGCGVLMVRKSASPSITTYECOTYERIAFSSRSKWIQFNSNEGNS 909
Qy 912 RGFCQPYVWV 922
Db 901 KGFPQVIVV 911

```


Matches 159; Conservative 72; Mismatches 224; Indels 78; Gaps 14;

Query CONTPTSYKSCSKRY-OGEGRQCDIDECGNGELNGCVCNHDCLNIPNYRCTCFDGFML 119

Db CLNTPGSYCKCKCGYLSTDKTCRYDLCATE-DHGCEQCLVNMIGSFVQCCSYTL 311

Query AHDCHNCLDVEDECLENNQCOHTCVNMGSYECCKEPFLSDNQHCTIRHSEEGUSMN 179

Db AEDGRCATMADYCASENHGCEHECVCNAESSSYLRCHEGFAINSKKTSKID---YCAS 367

Query KDHGSHICKEARGSVACBRCFELAKNQDC-LTCNHNGGGQHSCDDTADGPEC 237

Db SNHGCQHECVAQTSALCRCLKGFMNPDRKRRINCYCALNKPGCBEHCVCNEEGHYC 426

Query SCHEGYKMTDGRSCE-----REDTIVLETSNTSVWDG--DKEVKRLLNE 284

Db RCGRCYGNLDPNGKTCRSPVHCAQDGHGECQLCINTERSFVCCSSGELNDPLKT-SRAD 496

Query TCAVNNGCDRTCKTGTSGVHCCSPVGETLQDGKTKDIDECCQTRNGGCDHFCRN-UGS 344

Db YCJLSNHGCEYSCTNTDKSFAQCPEGHVLRSDGKTAKLDSCALGBHGCEHSCVSSEDS 546

Query FDGCGKKGKLUDEKSDPVDCC-SLRTCHSCINHPTACACRGYTL-VGTHG 402

Db FVCCPEGVILRDKTCCRDKDVGDNHGCHEUCLVNSGESVYKLGFRLAEDGSR CR 606

Query DTNCISINNGGQCVQCVNTGVSFECQCHPGYKJHWNKKCCE----- 444

Db RKAVCKS-QHGCSEMCVANGNSLICRGSEGFVIAEDGKCSCTEGPIDLVPGSKSL 666

Query VKGSLPT-SVSPRVLHCKSGGGDGFRLRGHSGIMHSDDVTITNSVT 492

Db GEENFETVKHFVQIIDSЛАVSPA-----ARVGLQIYSTQVRTEFT 768

Query FK---LNQGKCSKNAELFPEGGPALPERKHSVKEFSRVTNLTCGKOP 541

Db LRGSSAKEMKAVTHMKYNGKGSNTGLALKH-MPERSFTQVBGARPSTQVP 760

RESULT 2

Sequence 1: Application US/0889743

Patent No. 598163

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer J.

APPLICANT: Lal, Preeti

APPLICANT: Gorley, Neil C.

APPLICANT: Shah, Purvi

APPLICANT: Kaser, Mattie

ADDRESSEE: Incyte Pharmaceuticals, Inc.

NUMBER OF SEQUENCES: 4

TITLE OF INVENTION: HUMAN MATRILIN-3

CORRESPONDENCE ADDRESS:

STREET: 374 Porter Drive

CITY: Palo Alto

STATE: CA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Fastos for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/C8/897,443

FILING DATE: Filed herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/POCKET NUMBER: PF-0348 US

Query Match: 11.1%; Score 624.5; DB 5; Length: 1394;

Best Local Similarity 24.9%; Pred. No. 1.9e-36; Matches 240; Conservative 54; Mismatches 334; Indels 297; Gaps 48; SEQ ID NO: 36

Query DVDEAQRGDQD:ADLCONTPTSYKSCSKRYO-GEARQCDIDECGNGELNGCHD- 101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415 855-0555

TELEFAX: 415-845-4466

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

TYPE: amino acid

LENGTH: 638 amino acids

STRANDEDNESS: Single

TOPOLCY: linear

IMMEDIATE SOURCE:

LIBRARY: JTRSNKTO2

CLOUE: 68-719

J5-09-397-443-1

Query Match: 12.2%; Score 682.5; DB 2; Length 639;

Best Local Similarity 33.0%; Pred. No. 5e-41; Matches 163; Indels 59; Gaps 11;

Matches 159; Conservative 60; Mismatches 163; Indels 59; Gaps 11;

Query CONTPTSYKSCSKRY-OGEGRQCDIDECGNGELNGCVCNHDCLNIPNYRCTCFDGFML 119

Db CLNTPGSYCKCKCGYLSTDKTCRYDLCATE-DHGCEQCLVNMIGSFVQCCSYTL 311

Query AHDCHNCLDVEDECLENNQCOHTCVNMGSYECCKEPFLSDNQHCTIRHSEEGUSMN 179

Db AEDGRCATMADYCASENHGCEHECVCNAESSSYLRCHEGFAINSKKTSKID---YCAS 367

Query FDGCGKKGKLUDEKSDPVDCC-SLRTCHSCINHPTACACRGYTL-VGTHG 402

Db FVCCPEGVILRDKTCCRDKDVGDNHGCHEUCLVNSGESVYKLGFRLAEDGSR CR 606

Query DTNCISINNGGQCVQCVNTGVSFECQCHPGYKJHWNKKCCE----- 444

Db RKAVCKS-QHGCSEMCVANGNSLICRGSEGFVIAEDGKCSCTEGPIDLVPGSKSL 666

Query VKGSLPT-SVSPRVLHCKSGGGDGFRLRGHSGIMHSDDVTITNSVT 492

Db GEENFETVKHFVQIIDSЛАVSPA-----ARVGLQIYSTQVRTEFT 768

Query FK---LNQGKCSKNAELFPEGGPALPERKHSVKEFSRVTNLTCGKOP 541

Db LRGSSAKEMKAVTHMKYNGKGSNTGLALKH-MPERSFTQVBGARPSTQVP 760

RESULT 3

Sequence 1: Application US/0889743

Patent No. 598163

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer J.

APPLICANT: Lal, Preeti

APPLICANT: Gorley, Neil C.

APPLICANT: Shah, Purvi

APPLICANT: Kaser, Mattie

ADDRESSEE: Incyte Pharmaceuticals, Inc.

NUMBER OF SEQUENCES: 4

TITLE OF INVENTION: HUMAN MATRILIN-3

CORRESPONDENCE ADDRESS:

STREET: 374 Porter Drive

CITY: Palo Alto

STATE: CA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Fastos for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: J5/07/447, 343

FILING DATE: 27-FEB-1993

SEQ ID NO: 36

LENGTH: 1394

5:77-97-30

Query Match: 11.1%; Score 624.5; DB 5; Length: 1394;

Best Local Similarity 24.9%; Pred. No. 1.9e-36; Matches 240; Conservative 54; Mismatches 334; Indels 297; Gaps 48; SEQ ID NO: 36

APPLICANT: Yin, Wishan
 TITLE OF INVENTION: LATENT TGF β BINDING PROTEIN (LTBP)
 NUMBER OF SEQUENCES: 13
 ADDRESS: Williams, Morgan & Amerson:
 STREET: 7676 Hillmont, Suite 250
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77040

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/479,722B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/316,650
 FILING DATE: 30-SEP-1994

FILER APPLICATION DATA:

APPLICATION NUMBER: US 08/479,780
 FILING DATE: 18-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Russey, Shadley P.W.
 REGISTRATION NUMBER: 39,458
 REFERENCE/DOCKET NUMBER: 4100.000500/FJS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (713) 934-7300

INFORMATION FOR SEQ ID NC: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1253 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-479-722B-4

Query Match: 10.2%; Score 571; DB 3; Length: 1253;
 Best Local Similarity: 24.5%; Pred. No.: 1.e-32; Matches: 218; Conservative: 83; Mismatches: 296; Indels: 292; Gaps: 46;

Matches 218; Conservative: 83; Mismatches: 296; Indels: 292; Gaps: 46;

Query 1: 1156 VNTTSSYNCCTTHPRVLDASEKRRCIRPARBSNEQQEETDYGQJLQWEHLSDEYVCSRJVG 12:15
 Db 1: 682 ALKT-----PEAWMNSZCGGJLQCPGEVSADCFAPCOLCALTFQPSAGRATSCF--- 7:29
 Db 1: 1216 KQTTCTECCLYGEAWGY-QC-ATC-PLKSDSDAQLCNIPVIGRCP-YGRDMLNDPSE 12:7
 Query 2: 730 ---PGGGGLATKHDATSTDFCERVQCPGFFNTTTCIRCSPVG-----TYQEP 7:79
 Db 2: 1272 QYTPREADPYF7DQRFLNSPBLQE-EC-GIANGCENGRCVRQEGYTCDCLGJYHJLT 1:328
 Query 3: 780 GKNNCUSCPGNITIDRGSTM-ITOCKNRRGELGDF----TGYIES--PNY--PNY 8:29
 Db 3: 1329 AKMTCED----WNECDELNNRMISLCKNAKINTDGSYKLCCLPGVWPSDKPKVCTPNT 1:393
 Query 4: 830 PANTE 8:34
 Db 4: 1364 ALNLE 1:388

Query Match: 10.2%; Score 571; DB 3; Length: 1253;
 Best Local Similarity: 24.5%; Pred. No.: 1.e-32; Matches: 218; Conservative: 83; Mismatches: 296; Indels: 292; Gaps: 46;

Matches 218; Conservative: 83; Mismatches: 296; Indels: 292; Gaps: 46;

Query 5: 596 32 YPPGRGR---RAGRPQDVDEAQGQDCHADALCNPPTSKSCPKSYGYS---EGQCD 8:6
 Db 5: 519 LPSSRSAVEIAPIQVETDEERLNQNTC-GHSQCVPGPSDSVCHAGYRSHPCHRYCVD 5:96
 Query 6: 67 JDCGKNEINGGCVDCNLRNPYRCTFDGMM- AHDGHNCLDVEC---LENNGCQ 14:0
 Db 6: 597 KNCEEAERPCCGGKICKNTGSSYNCHNRYGRJRHGAGGRSVCVLINECAKPHUGBGG- 6:54
 Query 7: 41 HTCVNMGSYECCKEGGFFLSPNQHTCIRSBEGJSCMNKPGCHICKERPRGSVAC-E 1:99
 Db 7: 655 -CINFGHGXKCYCPSYRLKASRPICEDDE--CRDSTCPDGKCNP2-GSFKAIA 7:9
 Query 8: 200 CPGFEJAKNQRCDCILTNHNGGCHSCDSTDGPBCS---CHQYKMTDGRSCLER 2:55
 Db 9: 710 CGPGR-----SQGUGACR-----DNVCESEGTPCPCW----- 7:38
 Query 10: 256 EDVLEVYESNTSWVGDKRVKRRLIMETAVNNGCDRCKDTSTGVHNSCPVSGTLO 3:15
 Db 11: 739 -----CENLPGSYRC-----CAGGTR 7:77
 Query 12: 316 LDKTCXOIDEQTRN3GCDHCKNIVGSPGGCKKAFKLUDEKSQDVPDESLDRG- 7:74
 Db 13: 758 TGRLSCIDVDEGEAGKVQDQGIC-CNPQSPFCQCLISGHLSRSRSRSCEDIECDFBACT 8:17
 Query 14: 375 DHSCINHFPGTACACNRGYTLGFTC-GD-TNECSINNGG-QQVCVNTVSYECOCHPG 4:32

RESULT 4
 US-08-479-722B-4
 Sequence 4, Application US/08479722B
 Patent No. 6074840
 GENERAL INFORMATION:
 APPLICANT: Bonadio, Jeffrey

Db 1182 CVDIDECRELNQRGJLUCKSERCVNTSGSFRCVCKAGFTRSRPHGP 1226
 Cy 437 Sequence 3, Application US/08199780
 Db 3b Patient No. 5763416
 ; GENERAL INFORMATION:
 APPLICANT: Bonadio, Jeffrey
 APPLICANT: Goldstein, Steven A.
 TITLE OF INVENTION: Gene Transfer Into Bone Cells
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TX
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/199,780
 FILING DATE: 18-FEB-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: JMIC:002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 320-7220
 TELEFAX: (512) 474-7577
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1252 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-08-199-780-3
 Query Match: 9.9%; Score 553.5; DB 1; Length 1252;
 Best Local Similarity 24.6%; Pred. No. 2.1e-31;
 Matches 218; Conservative 82; Mismatches 300; Indels 285; Gaps 47;
 Qy 32 VPPGRGR--AAGPQEDVDEAQGLDDCHADALCQTPTSKSCPKPGQ- EGRQCB 86
 Db 538 LPPRSVAELIAPQTQETDECRNLQNC-GH3QCPGSPSYCHNAGYRSHPRHYCD 546
 Qy 87 IDECGNELLNGGCCWHDCLNIPSNYRCFGDGML-AFDGHNCIDVDEC---LZNGGCQ 140
 Db 597 VNECEAEPGCPGPKGICANTGSSYNCHNRGRYRHLWAGGGKRSVCD-NECAKPHLGDDG- 654
 Qy 141 HTCVUNGSYECCKEGFFLSDNQHCTIRSEEGLUSOMXDHGCHICKAEPROSVAC-E 199
 Db 655 -FCINPPGPHYKCNCPYGRKLKASRPPIECDIDE--CRDPSTCPDKCENKP-GSFKIA 703
 Qy 200 CRGGFELAKNQRCILTCNHNGQCQHSCTADGECSCSCHPKYKHTDGRSCBRETV 259
 Db 710 CQPGYR-----SQPGACR-----DNECS-----BTPC----- 734
 Qy 260 LEVTESNTTSVVDGDKERVKRLLMETCAVNINGGCDPTKDSTGYFRCSCRVGFTIQJDGK 319
 Db 735 -----SPOCKEK---LPGSFRCTCQGIRTRTGR 761
 Computer Readable Form:
 Medium Type: Floppy disk
 Computer: IBM PC compatible
 Operating System: PC-DOS/MS-DOS
 Software: PatentIn Release #1.0, Version #1.25
 Current Application Data:
 Application Number: US/08/3-6,650
 Filing Date: 30-SEP-1994
 Classification: 514
 Prior Application Data:
 Application Number: US 38/199,780
 Filing Date: 30-SEP-1994
 Attorney/Agent Information:
 Name: Parker, David L.
 Registration Number: 32,165
 Reference/Docket Number: JMIC:002
 Telecommunication Information:
 Telephone: (512) 418-3600

Db 822 INTNGSYRCUCLPGLHRGLVGGRKCKKDIDCSQPSLCLPHACENLQGSYVCVCEBPLT 881
 Cy 437 MKKDCVEVKSLPLTSVSPRVLHCKSGGGDFJURCHSGIHLSDDVTRISUTKLN 496
 Db 3b Sequence 3, Application US/08199780
 ; GENERAL INFORMATION:
 APPLICANT: Goldstein, Steven A.
 APPLICANT: Bonadio, Jeffrey
 TITLE OF INVENTION: Gene Transfer Into Bone Cells
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TX
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/199,780
 FILING DATE: 18-FEB-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: JMIC:002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3600

Db 982 QDQHGBEEVE-----OPHMKKE---CYLNFDTWFCDS 1-VLATNT- 920
 Cy 497 ECKCSL----KNAELFPEGIRPALEPEKHSSVKEFRRYVNLTCSGCKQVPGAPRSPSTPK 551
 Db 921 ECOCSTGAGWDHCE-YPCPVYSS-AEFHSLVDPGR-----LHSQO----- 962
 Qy 552 EMFIVFRELEINNOKEVITASDLSCLVKRERKARERKAIRHREOZHOLSGMND 611
 Db 963 -----REIC-----HELI-CI-----PAHDIDECILPG---- 981
 Qy 612 VAKKPPRTSERORESCVGQGCHAENCVS-----CRAKYDGCARECI----- 655
 Db 982 -----REIC-----KEGKVNSQBGYECYCKQYDNQJEDVNECDE 123
 Qy 656 -LCPNGTPTONIEEGQMTCEPCH---RPGNSALKTEPAWNSE---CGG-JQPGBSYAD 706
 Db 1024 SNRNGYGEN-TWLPLPACTPAAEPAQACL-IPEWSTPQRDVKCAGAGE-BRTAC 1079
 Qy 707 GAFRCQCALCF-----QEAGRSCFC-----GGGLATKHGATSEDC----- 748
 Db 108C VMQPWAGA-----FDCCORPLRG-TOCRPDRPPRGTSQCBTSQSBSNSFWDTSPLLGK 1137
 Qy 749 -----TRVQCPGPHFTNTTHRCIRPVGTQPEFEKRANNVSCPGN-----T 791
 Db 1128 SPDDEOSSEESDEC-----Rvsgvprvrgavc-EGRCGQCLASRAR 1182
 Qy 792 TIFPGSTNTI---CNRGZGELGF-----TCVIES PNYP 826
 Db 1133 CVDIDECRELNQGLLCKSERCVNTSGSFRCVCKAGFTRSRPHGP 1227

RESULT 7
 US-08-316-650-3
 Sequence 3, Application US/08316650
 Patient No. 5942496
 ; GENERAL INFORMATION:
 APPLICANT: Goldstein, Steven A.
 APPLICANT: Roessler, Blake J.
 APPLICANT: Goldstein, Steven A.
 APPLICANT: Goldstein, Steven A.
 Title of Invention: METHODS AND COMPOSITIONS
 Title of Invention: FOR STIMULATING BONE CELLS
 Number of Sequences: 25
 Correspondence Address:
 Addressee: Arnold, White & Durkee
 Street: P.C. Box 4433
 City: Houston
 State: Texas
 Country: USA
 Zip: 77210
 Computer Readable Form:
 Medium Type: Floppy disk
 Computer: IBM PC compatible
 Operating System: PC-DOS/MS-DOS
 Software: PatentIn Release #1.0, Version #1.25
 Current Application Data:
 Application Number: US/08/3-6,650
 Filing Date: 30-SEP-1994
 Classification: 514
 Prior Application Data:
 Application Number: US 38/199,780
 Filing Date: 30-SEP-1994
 Attorney/Agent Information:
 Name: Parker, David L.
 Registration Number: 32,165
 Reference/Docket Number: JMIC:002
 Telecommunication Information:
 Telephone: (512) 418-3600

Db 822 INTNGSYRCUCLPGLHRGLVGGRKCKKDIDCSQPSLCLPHACENLQGSYVCVCEBPLT 881
 Cy 437 MKKDCVEVKSLPLTSVSPRVLHCKSGGGDFJURCHSGIHLSDDVTRISUTKLN 496
 Db 3b Sequence 3, Application US/08199780
 ; GENERAL INFORMATION:
 APPLICANT: Goldstein, Steven A.
 APPLICANT: Bonadio, Jeffrey
 TITLE OF INVENTION: Gene Transfer Into Bone Cells
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TX
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/199,780
 FILING DATE: 18-FEB-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: JMIC:002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3600

Db 1344 CG-----PETS PGWECVDNECELMVAGD ALCEN VEGSFJC 1382 Qy

Cy 588 AIZTRKAVHREQFH- QUSGMNIDVAKK PERTSERQAS-----CG- 628 Db 1383 LCAADSEVDAEEHCRRPRAVG---AQPIEVRTENQAPSILRIMCYSEHNGPPSQZ 1138 Qy 629 VGOQHAEAOCVSFRAGTYDGA RERC-LCPNGT FNEEGQMTC EPPR-----PG 673 Db 1439 LGQNSTQREC-CCTQGARW GKA-----CAPPS ESEVF SQICPS 1477 Qy 679 NSGALKTPRAW-----NMSEC---GGLCQFGEYSA DGFAPCCALGTFOEASRT 726 Db 1478 GOGYI PVEGA WTFOGTMVTDADBCVLFGP ALCON GROS -NIVPGYIC----- 1524 Qy 727 SCFFC3GGLATKHQGATSPQDCBTRVQCS PSHYNTTHR C-----IRCPVQYQPE 778 Db 1525 -----CNPQHYDASSKQCOHDQDIA ENCEVQ 1557 Qy 779 FGKNCVSCPNNTIDFGSTNITQCKURRGEBLGD F-----TGYIESPN 824 Db 1558 EGSPHCL CNPPLTUDLSQ-----RCVNTTSSTE -DFPDHDHM DICWKVWINDVSCP 1610 Qy 825 YPGNYPANTEC 835 Db 1611 LRGHTTVTEC 1621

RESULT 10

; Sequence 2, Application US/09230652A

; Patent No. 653775

; GENERAL INFORMATION:

; APPLICANT: Tournier-Lasserre, Elisabeth

; APPLICANT: Bousser, Marie-Germaine

; APPLICANT: Bach, Jean-Francois

; TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND TREATMENT

; FILE REFERENCE: 03715.0048-00000

; CURRENT APPLICATION NUMBER: US/09/230,652A

; CURRENT FILING DATE: 1999-05-17

; EARLIER APPLICATION NUMBER: FR 96 09733

; EARLIER FILING DATE: 1996-08-01

; EARLIER APPLICATION NUMBER: FR 97 04680

; EARLIER FILING DATE: 1997-04-16

; EARLIER APPLICATION NUMBER: PCT/FR97/02433

; EARLIER FILING DATE: 1997-07-31

; NUMBER OF SEQ ID NOS: 163

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 2321

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: human ADNC No. 6537775ch 3

US-09-230-652-2

Query Match: 9.4%; Score 526.5; DB 4; Length 2321; Best Local Simil.arity 24.2%; Pred. No. 3; Del-29; Matches 232; Conservative 75; Mismatches 303; Indexes 347; Gaps 58.

Qy 33 PFG-RGRAGPQEDVDEAQGIDDDCHAQDAACTT PYSKCKPQG----- 79 Db 147 PPGYQGRSC-RSDVDECRVG-BPRHGCTLNTPSFRCQCPAGYTGIPLENPAVCPA 203 Qy 80 -----BRCOE-DIDGCNE-LNSGVHVNININGRYC 111 Db 204 SPCCRGGT CROSSEL TDPACI PGPFGEGCENVNDP CPHRCINGGT CVDGN----- 260 Qy 112 TCGFPMALHCGNCL-DVDECLENNGQOH-T CVMGMYECCKEGF---FISDN- 163 Db 261 QCPPEW---TGGCTEDVDEQPNACHNGCTCFTTUGGSCVYNGESSIOND 316

Qy 164 ---QHTCH-----RSEEGLUSCMKDHGS SHIKE-----APRSVA 197 Db 317 DATAVCEHGATC HDVASPFGACPMKGTG LICHEDACUSNPCHEDA-COTNPNGRAI 376 Db 138 CGRPGE LAKQDRCJU TGNGSACOH - SCDTADGSECSCHQY -- KMITDGRC 252 Db 377 CTPPGFTGAGDQD-VDECSIGANPCEHLGR CUVNQGSPLOC CRRGTYGPRCETDVNEC 435 Qy 253 J---ERBTIV-EVTENNTSUDGKVRKVRLLMTCAN- NGCDTCKTGSTGH 305 Db 436 iNSPCRQIATC DRIQFTCIMAGITGTYCWD-DECOSPCNGG--VKORVNGS 492 Qy 306 CSAPVGF TLQDGKTC-CIDBCQ----RNQGCDHICKNIVGSPQGCKKGKFLJDEK 360 Db 493 CTCPGFS---GSTCQLDNVCAS-PCRNHA-- KQVDPDQYECRCAEGEGFLCD- 543 Qy 361 SCOPDVS CDSLRTCDHS-CINHPGTFACACRGRGYTYGFTC----- 401 Qy 544 -RNDDCSPD-PCHHGRCVGDIAS-PSCACRGPY---GTRCESQDEC RSQPCRHGGKC 597 Db 402 -----GDTN-EC51NGCOC-----WCVNTGSYCQCHGKLYLWNKKC 441 Db 598 LENDVRLKLCRFSGTTIVNCVNNIDCASNPTGFGVODRINRYD-VCCG----- 649 Qy 442 CIVVKGQIPTSVSPRVS LHC KSGG---GDOCF ARCHSGTHSSCVTIRT SVFKLNEG 498 Db 682 -----LCPGS LPLCJLSSPHRCHE-----PCSHGICXDPAGG----- 715 Db 630 --TGPLCNVNEIN CASSSP CGBGGSCV DGENF-RC----- 681 Qy 555 IVEFELETNQEVTA SCDSLSCV KRTEK RLKAIR TLRK AVHRE QFHOLQSMNLDVAK 614 Db 716 -----FRCV----- 724 Qy 615 KPPRTS RQAES CGVQG HAGA ENQCVS---CRA G-----TVDGARERC-LCPNGT FNEEGQM- 669 Db 725 SGPRS-----OSLAR DACESOP C PRCAGT G-----SSDG MGFHC-TCPG YV-QSRQCELL 772 Qy 670 -ICEPCP-----RPGN SALKIPEA NMNS ECGL CQZ P23 EYSA DG-FAR CQCAL GTF 719 Db 773 SPCTPN CEBHSGRCESAPGC-PVCSOCPQGNG---SPRC COVND ECA G PAGC- PHGIC 826 Qy 720 QPEAGR ITSCOFFGGJATKHQGATSPQDCETRVQCS PGP HFNNT-THCIRC P VGYQPE 779 Db 827 TNLAGTSSC-TCHGG---YPGSCODIN--DODPNCING-----GSCQDGY 868 Qy 780 GRKNCVSCPNNTIDFGSTNITQCKURRGEBJ3DPFGYIES-----RNYPG 927 Db 869 GSFCSCSLPG-----FAGP RCRD VDECLS N P C GP- GTC T D HVA SFTCT CPPG YGG 918

RESULT 11

; Sequence 19, Application US/08185432

; Patent No. 5750652

; GENERAL INFORMATION:

; APPLICANT: Artavanis-Tsakonas, Spyridon

; APPLICANT: Biseau, Isabelle

; APPLICANT: Ederich, Robert J.

; APPLICANT: Xu, Tian

; APPLICANT: Yamamoto, Kenji

; TITLE OF INVENTION: DELIX PROTEINS, NUCLEIC ACIDS, AND RELATED METHODS AND COMPOSITIONS

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDWONDS

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DCSYMS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/185,432
 FILING DATE: 21-JAN-1994
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Vischeck, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NC: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2523 amino acids
 TYPE: protein
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 JS-08-185-432-18

Query Match Score 9.9%; Score 501.5; DB 1; Length 2523;
 Best Local Similarity 23.3%; Pred No. 2.6e-27;
 Matches 233; Conservative 99; Mismatches 300; Indels 383; Gaps 65;

QY 33 PPGRRRAAGG---QEVDDEAQGLDCHAHALCQNTPTSKCSCKRPGYQEGEGRCDE-DDE 89
 Db 400 PPG---YTGPACPNVNDPDECSLIGANPCEHGRRTNTGSFOCNCPOGYACP--REBIDNE 454
 QY 9C CGNELNGGCDP---CLNPWNYCTCFDGFMMLAHPGHNC-LDVDECLENGGCHT-CV 144
 Db 455 C---LSNPCCNDSTGLDQIGEFCOMPY---EGLYCETNTIECASN--PCC.HNGKCI 505
 QY 145 NMGSVECCKEGGFLSDNQHITCHRSEE---GUSMNMKHGCHICKEAPRG-- 194
 Db 506 DKINEFRCDPFTG---SGNLQHDFECTSTOCKANGAKLDGNSYRCQCTEGFTGH 56:
 QY 195 ---SVACERPGF---JAK---NQDCC--- 213
 Db 562 QEDDNECIPDPCHYCTKDGIAATCLOCRPGYIGRLCENDINCLSKPCLLNGQCTDRE 621
 QY 214 ---ILTNHNGG---QHSCDTA-----DEPECSCHPQYKHTDGRSLCR 255
 Db 622 NGYICCPKPTGTGNCETKDCCASNLCDNGKCIDKIDSYCETCPY---TGSIC--- 674
 QY 256 EDTVLUVTESMTSVDGDKRKVRKRLMTCAVNGGCR-----TCK 298
 Db 675 ---NINNECDNSPQRNGG-TCKQDINGTCVCGDGHOMCLSEVNENNSNPCNGH 730
 QY 299 DSTGPHCSCPVGFLDQGCKTKD---DECOTR---NGGDHCCKNNGSFDCGKGF 353
 Db 731 DGVNGYKCDBEAGWS---QSNIC---JINNECESNPOMNGT---CKXW-GAYIYCAGF 782
 QY 354 KLTDEKSCQ-CDECSLORTCDHS-----CINHP--- 382
 Db 782 ---SAPNCQNNINNECNSNPOLNFGTCIDWAGYKCKMLPYGKICEAVLAPCAGSPCK 838
 QY 383 -----GTFACACNRYGTYLGTHGECENECS:NGGQ---VCNTVTSYECQ 429
 Db 839 NQGRCKESEDEPTESECPPGM---QGOTSEIIMEXCV-NRPERGATCONTNGSYKNC 894
 QY 430 HPGYKJ-HWKUDC-WVKHJLPTSVPRVLSLHGKGSSGGDGFRLCHSGHILSSDTTR 498
 Db 895 XGY---TGRNCMIDCQDQNP-----CHNG----- 918
 QY 489 TSVTPLKNECKSLNAEL---PEGRLPALPEKHSVSKESPRVNLTCSSGKV--- 540
 Db 919 -----GSCS-DGINMFNCPCAGFRG---PKCEDINEC---ASNPKRNGANCTDCV 963
 QY 541 -----PGAGR-STPKEMFITVEFELETNQEVUTASCLDSC-VKR-EKRLKAIRL 592

RESULT 13
 Sequence 18, Application US/08185432
 Patient No. 575052
 GENERAL INFORMATION:
 APPLICANT: Aravanis-Tsakonas, Spyridon
 APPLICANT: Busseau, Isabelle
 APPLICANT: Diederich, Robert J.
 APPLICANT: Xu, Tian
 APPLICANT: Matsuno, Kenji
 APPLICANT: DITEX PROTEINS, NUCLEAR ACIDS, AND COMPOSITIONS
 TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:

Db 964 NSYTCTCQPRGSG-----ZHCESNTPDCTES---SCFNGGT----- 996
 QY 593 RKAVERREQFHQLQSGNLDWAKKPPRTSERTQAEASCGVGCHAENQCVS--CRAGTYDGA 65C
 Db 997 -----CIGINTFCQCPGFGTGSYCO-----HDNECSDKPCLNG----- 1032
 QY 651 RERCICLUPNSTFONEPGOMTCPEPDRPGNSGALKTPPEAW-NMSBC--GGICOPGEYSADG 707
 Db 1033 -----GTCQDSYSTYKC-TCPQGYTGUNCONLVRWCDSFCKNGKC---WQTNH 1078
 QY 708 FAPCQCALG---TSPEARTRSCPCGGGLATKEQGATSFQD-ETRVCSPCHFYNTH 765
 Db 1079 FYRCE-CKSGMTGVTGVDVPSVC---EVAKOQGVDVHLCRNSGMVD---TGTH 1128
 QY 766 RCIRCPVGTQPERFSGKNVCSCP-----GNTTDFDG-----SNTIQC 804
 Db 1129 FC-RQAG---YGSTCSEQEDCSNPONGATCPCDLYGSVCEVAGYHGNCBEEINEC 1146
 QY 805 KNRRC--GGELGDFGYIESPNYPNPYANT-----ECT-----WTNP 841
 Db 1187 LSHPCONGSTIDLN---TYKSCSPROTCQGVHCEINVDCDTPFYDSPTJEP 1235
 RESULT 14
 JS-08-899-232-3
 ; Sequence 3, Application US/08899232
 ; Patent No. 6436650
 ; GENERAL INFORMATION:
 ; APPLICANT: Aravanian-Tsakonas, Spyridon
 ; APPLICANT: Qi, Huijie
 ; TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
 ; FILE REFERENCE: 7326-046
 ; CURRENT APPLICATION NUMBER: US/08-899,232
 ; CURRENT FILING DATE: 1997-07-23
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2523
 ; TYPE: PRM
 ; ORGANISM: Xeropus sp.
 ; JS-08-899-232-3
 Query match 8.9%; Score 501.5; DB 4; Length: 2523;
 Best Local Similarity 23.0%; Pred. No. 2.6-27;
 Matches 233; Conservative 98; Mismatches 300; Idents 383; Gaps 65;
 CQ 33 PPGRGRGAAGP---QEDVDECAAGLQDCHADALCQNTPSYKSCSCPGYQSEGRCOE-DIDE 99
 Db 400 PPG---YTPGACNDVDECUSGANPCEHGRCTNLGSQCNCFQGYAGP---RECIDNE 454
 QY 90 CGNE-LNGGCVID---CLNPENSYRTCTCFGMLAHGHNCLDVECLENNGGCQT-CV 144
 Db 455 C---LNPCCQNDSTCUDQIQSFQCIQPGV---EGLYBTNTDECASH---PCHNGKCI 505
 QY 145 NVNGSYECCKEGFFLSQDQHTCIRSEE-----GLSMNKDHGSHICKAEPRG--- 194
 Db 506 DKINEFRDCDCTGEG---SGNLQHDFDECSTPCNGAKOLDGPNSYTQCTEGFTSRH 561
 QY 195 -----SVAECPGFE-----LAK---NQDCC--- 213
 Db 562 CBQDNECIPPCPHYCTKODGIATFCCLCRPGYGRLCNDJNFCLSKCLNGSQCTBRE 621
 QY 214 ---ILTCNHGNGG---QHSCDDTA-----DGBECSTPHOYKHTDGRSCER 255
 Db 622 NGYICCPKGTMGNGCSTKDDCAASLNUCDNGKCKDIDKVECT-EPGY---TQKLC--- 674
 QY 256 EDTVLEVTESNTSNTWDGDKRVRKRLMETCAGVNGGCR-----TCK 298
 Db 675 ---NINNECDNSNCPNRNGG---TCKDOINGFTCVCPODYHDMCLESNECNSPCIHACH 730
 QY 299 DTSTGHSCSCVGFITQDGKTKD---DGCCTR---NGCDHFFKN-VGSDFGCKGF 353

Db 731 DNGNSYKCDCAGNS---GSNC-DINNECESNPMNGT--CKDTGAVIETCKAGF 782
 QY 354 KLLTDTEKSCO-DYBECSDLRDCDS-----CINHP--- 382
 Db 783 ---SGPNCONINNECSSNPCLNHGTCIDDVAGYKCNMJPYGAICEAVLPCAGSPCK 838
 QY 383 -----GTPCACNRGYTLGFGHGDNECSINNGCQ---UCVNNGSYECQ 429
 Db 839 NGCRCKESEDFETSCCECPGPW---CGOTCEIDMNNECV--NR-PRGRNGATQNTNGSYKCNC 894
 QY 430 HPGYKUHNKRD-VEVKGLAPTSVSPRVLHCKSKGGGCFURCHSAHJSDVTIR 488
 Db 895 RPGY---TGNCEMDIDCQPNP-----CHNG----- 918
 QY 489 TSUTVFLNEBGRCSLNAELP---PEGLRPAPRKISSKESPRYVNLTQCSQH---- 540
 Db 919 -----GSCS-3GIMMFNCNPAGFGR---PKCEDDINEC---ASNPKCKGANCNTDCV 963
 QY 541 -----RHPGRPSTPKEWTFEFETNQKEVTSACDSLSPKRVKRELRKAIHL 592
 Db 964 NSYTCTCQPRGSG-----HDNECSDKPCLNG----- 996
 QY 997 -----CIGINTFCQCPGFGTGSYCO-----HDNECSDKPCLNG----- 1032
 Db 651 RERCICLUPNSTFONEPGOMTCPEPDRPGNSGALKTPPEAW-NMSBC--GGICOPGEYSADG 707
 Db 1033 -----GTCQDSYSTYKC-TCPQGYTGUNCONLVRWCDSFCKNGKC---WQTNH 1128
 QY 766 RCIRCPVGTQPERFSGKNVCSCP-----GNTTDFDG-----SNTIQC 804
 Db 1129 FC-RQAG---YGSTCSEQEDCSNPONGATCPCDLYGSVCEVAGYHGNCBEEINEC 1146
 QY 805 KNRRC--GGELGDFGYIESPNYPNPYANT-----ECT-----WTNP 841
 Db 1187 LSHPCONGSTIDLN---TYKSCSPROTCQGVHCEINVDCDTPFYDSPTJEP 1235
 RESULT 15
 JS-08-899-230A-20
 ; Sequence 20, Application US/08899230A
 ; Patent No. 586558
 ; GENERAL INFORMATION:
 ; APPLICANT: Aravanian-Tsakonas, S. et al.
 ; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 ; TITLE OF INVENTION: And Compositions Based On No. 5786193Ch Proteins And
 ; TITLE OF INVENTION: Nucleic Acids
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds, The Americas
 ; STREET: 115 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/393,590A
 ; FILING DATE: 25-JUN-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: M. Strock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-015

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 8688479741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEC ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2556 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 JS-08-093-590A-20

Query Match 8.8%; Score 491.5; DB 1; Length 2556;
 Best Local Similarity 23.7%; Pred. No. 1.4e-26; Mismatches 339; Index 337; Gaps 51;
 Matches 238; Conservative 90;

Qy	41 GP--QEDVSEGAGLDDCHADLACNTPTSYKSCK-----	Db	172 GPRQDPNECCGQKPRILRRHGGTCHNEVSYRVCRAHTGNCERPVYPSPSPSQNGG 231
Qy	75 -----PSVQGEGEROEDIDEC-GNEL-NGCVAQDCJN:PGNYRCTCPDFE 117	Db	232 TCRRTGDSVTHECAGLPGTQNC-E-NIDPCRNCEKNGGAGTGIN-----TWCRCPPW 267
Qy	118 MIAHDGHVCL-DYDTCLENNNGSCH--TCAVNGSYGCCKEGFF---LSDN----QET 165	Db	289 ---TGQYCTEDVDECOLMPNACONGGTCHNTTGGNCVUVCNGWTGDCSEN:DCASAA 343
Qy	167 CIH-----RSEPLSCMNDKHCSSHKKEAP-----RGVACCRPG 203	Db	344 CPHATCHDRVAVSFYCSCPHGRGLLCHINDACTISNPNEGNSNCDTPVNGKAICTCPSPG 403
Qy	204 FELAKNORCIL:FNHNGNGCQHS--CDDTADGFECSCPHQY---KMTDGSCS-----E 254	Db	404 YTPBACSGD-VDESLGANPCEAGKINTLSEFCQOLGGYSPREIDNECVSNQQ 462
Qy	255 REDTVVEVTESNTTWSVPPDKRVRRLMETCA---VNNGGCDRTCKDTSTGVHSCPV 310	Db	463 NDATCLDQIGEFQOMCYFGEGYHVCEVNTECASSPLHNG---RCLDKINBFCBCFT 518
Qy	311 GFTQLDAGKCK-DIDEQT--RNGGDHFCKN:VSPFCGCKKGKPLLT3EKCQDV 366	Db	519 GFT---GHLCOVYDECASTPCKNGA--KCJDGPNTYTCVTEGY--TGTHCEWD 568
Qy	367 ECSVDRTOHCSINCINPGRIFACAGCRNGTYLGFTIC-GTNECSIN---NGGQQ-----416	Db	569 ECPDPCHGSGCKGQVATTCLOPREGYCH--HECIMINESSQPRLRGGCOPDNAY 625
Qy	417 -----VCNTVPSYCOCHROYKHMKKZCCEVKG 448	Db	626 LCFLUKG:TOPNCIIDLDCASSCDSGTCJDKIDSYCACERFGYT3SMCNVIDEAG- 684
Qy	449 LPTVSPrVLHCKSG---GGCFURCHSGHLSCLWTT-RTSVFKLNEKGCLN 504	Db	685 -----NP-----CHNGSGEDGNGFTCRPEGH--SPTCLS-----EVNB--CN--- 721
Qy	505 AELPPEGURRALPEKKISSKESFVVNUCTSSGKQVGPGRISTPKEMFIVPELETN 564	Db	722 -----SNCFVHGACRDSDJWYKDCCD-----PWSG----- 747
Qy	565 QKEVTASCDLS-----CIVKRTKRLKAITTRKAVRECFH:QLSGNMLDVAKKP 616	Db	788 ---CQTNINEC-----ASNPLN--KGTCIDDAVAGYKCNCLJPYTA-----TCEV 828
Qy	748 -----TNCINNNCESNCVNGGTCIKDMTSGI---VCTCREGF---SGP----- 787	Db	673 ---PC-PRO-GNSGALKTEBWNNSC-----GGCQPSYSADGFAPCQLCALGTF 719
Qy	829 VLAAPAPSPORNGBCRQSBDYEFSCVPTAGKQ-CEV-DINEV:SPERHGA--- 883	Db	

Search completed: October 22, 2003, 11:31:49
 Job time : 39 secs

Qy	720 QPEAGR-TSCPPCGG-LATEHQGATSPQOCETRV-OCCPQHFT-----NTHRCIR 769	Db	884 -----SCONTGSGYRCGQAGYSGRNCETDIDCRPN:CHNGSCTDGINAFCDLP 936
Qy	770 CPVGTGQPERGKNCVSCP---SNTTD-----FQG---SNTICKNRCC-- 809	Db	937 CPFRGTFCEE-DINBCASCPCRNGANCTDCVSYCTCPAGFSSGIHCENNTPOCTESSCFN 995
Qy	810 GSE---LGPTGVIESPNNPQNPANTECTWTINPPKRI:LI 849	Db	936 GGTGCGGINSFT-GLCPGFTGSY-----QHVNNECSRPLS 1032

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OX protein - protein search, using sw model

Run on: October 22, 2003, 11:24:13 ; Search time 121 Seconds

Sequence: 1 MCVAGRNRPQAAWAVLILL.....PRSFRLIRRSRVSRLRPYK 999
2130.533 Million cell updates/sec

Title: US-09-747-371-2
Perfect score: 5605
Sequence: 1 MCVAGRNRPQAAWAVLILL.....PRSFRLIRRSRVSRLRPYK 999
Scoring table: BloSUM62
Gapcp 10.0 , Gapext 0.5

Searched: 830525 seqs., 258052604 residues

Total number of hits satisfying chosen parameters: 830525
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match: 0%
Listed first 45 summaries

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match length	DB ID	Description
1	5605	100.0	999	4 QPNQ36	Qpnq36 homo sapien
2	5056	90.0	997	11 Q9JJS0	Q9jjs0 mus musculu
3	3662.5	65.3	988	4 Q8IWY4	Q8iwy4 homo sapien
4	3640.5	65.0	992	11 Q8C9Q4	Q8c9q4 mus musculu
5	3397	60.6	993	4 Q8IX30	Q8ix30 homo sapien
6	3262.5	58.2	961	11 Q9EQC6	Q9eqc6 mus musculu
7	2916	52.0	880	4 Q8NAU9	Q8nau9 homo sapien
8	1738.5	31.0	581	4 Q8NVA8	Q8nva8 homo sapien
9	875	15.6	1664	5 Q9TQ02	Q9tq02 cencorhabdi
10	84.5	15.0	1574	11 Q88281	Q88281 rattus norvegicus
11	775.5	13.8	1409	5 Q9VB89	Q9vb89 drosophila
12	737.5	13.2	1246	4 C15095	075cs hom sapien
13	720	12.8	2906	11 Q9WCH9	Q9wch9 rattus norvegicus
14	710.5	12.7	2899	4 Q96JP8	Q96jp8 hom sapien
15	695	12.4	810	11 Q8C435	Q8c435 mus musculu
16	695	12.4	11 Q99K64	Q99k64 mus musculu	

17	695	12.4	956	11 Q8R542	Q8r542 mus musculu
18	695	12.4	1289	5 Q8SSS3	Q8sss3 dictyostellum
19	694	12.4	2189	5 Q9B105	Q9b105 emerita tenella
20	693	12.4	3857	1 Q88840	Q88840 mus musculu
21	686.5	12.2	937	4 Q96FT5	Q96ft5 homo sapien
22	686.5	12.2	1016	4 Q8NDE6	Q8nde6 homo sapien
23	675	12.0	2872	11 Q9WUFB	Q9wufb rattus norvegicus
24	644.5	11.7	161	4 Q9Y559	Q9y559 homo sapien
25	640	11.4	708	13 Q87363	Q87363 gallus gallus
26	610.5	10.9	1062	1 Q60789	Q60789 mus musculu
27	610	10.9	1398	13 Q8AXM6	Q8axm6 xenopus laevis
28	607	10.8	1399	13 Q8JF24	Q8jf24 xenopus laevis
29	601	10.7	1389	11 Q8CG18	Q8cg18 mus musculu
30	601	10.7	1713	11 Q8CG39	Q8cg39 mus musculu
31	598	10.7	1713	11 Q88349	Q88349 mus musculu
32	597.5	10.7	1963	6 Q28319	Q28019 bos taurus
33	587	10.5	1382	4 Q9H7K2	Q9h7k2 homo sapien
34	574	10.5	1972	4 Q9WPL1	Q9wpl1 oikopleura dioica
35	563.5	10.4	1764	11 Q35806	Q35806 ratus norvegicus
36	575	10.3	893	6 Q8MK0	Q8mk0 cercopithecus aethiops
37	575	10.3	733	5 Q25678	Q25678 podocoryne
38	575	10.3	1554	5 Q8NGV4	Q8ngv4 drosophila melanogaster
39	575	10.3	3396	5 Q8V555	Q8v555 drosophila melanogaster
40	574	10.2	1174	11 Q99E58	Q99e58 mus musculu
41	571	10.2	1253	11 Q67810	Q67810 mus musculu
42	570.5	10.2	1821	4 Q14767	Q14767 homo sapien
43	567	10.1	638	4 Q8NB66	Q8nb66 homo sapien
44	567	10.1	685	14 Q922KB	Q922kb8 mus musculu
45	566.5	10.1	746	4 Q96HBS	Q96hbs homo sapien

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

ALIGNMENTS

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Database : SPTRIMBLI 23;*
1: sp_archea;*
2: sp_bacteria;*
3: sp_fungi;*
4: sp_human;*
5: sp_invertebrate;*
6: sp_mammal;*
7: sp_mhc;*
8: sp_organelle;*
9: sp_phage;*
10: sp_plant;*
11: sp_Rodent;*
12: sp_virus;*
13: sp_vertebrate;*
14: sp_unclassified;*
15: sp_virus;*
16: sp_bacteriap;*
17: sp_archae;*

?pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db	479 SSDTIVTURSVTKNEKGCSLSQKAKSPEGRPRPDEHSYKESQYANTICSTCOKV	538
Qy	541 PGAPRGPSPKEMITVFELETOKEFTASCDSLCTVKTERLRKAIRTRKARHREQ	600
Db	539 PGALGRLNAPKEMITVFELEYEKEVFTASCNLSCVTKTERLRKAIRTRKARHREQ	598
Qy	601 PHOLSGMNLDAVKPKRPTSERQASCGGQHANQUSCRAGTYDARGERCILCPNG	660
Db	599 FHQLSGMNLDAKPTPSRSVQSGHQETCQVGQHEESQVSACRAGTYDQGSOBCILCPNG	658
Qy	661 TFQNEBEGOMCERPFRPGNSGALKTPBAWNMSZCGLGICOPGEYSAQDAPCOLCAGTFQ	720
Db	659 TFPNEBEGOVTCEPGRPENGLSKISEAWNVDGGLQUPGESSANGFAPCQDCALGTFQ	718
Qy	721 PEARGTSCPCCGGAATKHQGATSFQDETRVQSCPGFYTNTTHRRCRCPVSYTQPFQ	780
Db	719 PDVGRTSCLSCGGSSGSIPLTHLGATSPQCFTRVQSCPGFYTNTTHRRCRCPVSYTQPFQ	778
Qy	781 KANQSVCPNTTDPDGSTNITQCRNRKRGEGLGDFQYIESPNYGNYPANSECTWNTIN	940
Db	779 KANQSVCPNTTDPEGNSMITQCRNRKRGEGLGDFQYIESPNYGNYPANSECTWNTIN	938
Qy	841 PPKRKRILIVWPEFLPTEDDCDYLVKRTSSNSVNTYECOTYERPIAFTSRSLW	900
Db	839 PPKRKRILIVWPEFLPTEDDCDYLVKRTSSNSVNTYECOTYERPIAFTSRSLW	898
Qy	901 IOKSNEGNSARGFQNVPTDQYDQELIEDVIRGRLYASENHQELIKDKLKLALPFV	960
Db	899 IOKSNEGNSAROFQPVYDQYDQELIEDVIRGRLYASENHQELIKDKLKLALPFV	958
Qy	961 JAHFONYFKYTAGSREMPRSFIRLLRSKVSFRFLRKY	999
Db	959 LAHPQNYFKYTAQESREMPRSFIRLLRSKVSFRFLRKY	997
RESULT 3		
Q8WY4	PRELIMINARY; PRT; 988 AA.	
AC	QIWW4;	
DT	01-MAR-2003 (TREMBlre. 23, Created)	
DT	01-MAR-2003 (TREMBlre. 23, last sequence update)	
DE	Signal peptide-CUB-EGF-like domain containing protein 1.	
OS	Homo sapiens (Human)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
OX	NCBI_TAXID=9606;	
RN	{1} SEQUENCE FROM N.A.	
RX	MEDLINE:22336413; PubMed=12270931;	
RA	Yang, R.B.; Ng, C.K.D.; Wasserman, S.M.; Colman, S.D.; Shenoy, S.,	
RA	Mehraban, F.; Komives, L.G.; Tomlinson, J.E.; Topper, J.N.;	
RT	"Identification of a Novel Family of Cell-surface Proteins Expressed in Human Vascular Endothelium."	
RT	J. Biol. Chem. 277:46364-46373 (2002).	
DR	EMBL; ARS25269; AAN77331; -.	
SQ	SEQUENCE 988 AA; 107901 NM; 1BF57BEF778CC9C1 CRC64;	
Query Match	65.3%; Score 3662.5; DB 4; Length 988; Best Local Similarity 62.2%; Pred. No. 2.96-317; Matches 652; Conservative 113; Mismatches 164; Indels 119; Gaps 24;	
Qy	:0 GAA---WAVLILMLPPLLALLAGVPPGRGAAQ----PQEDVDEAQGLDDEHAL	62
Db	2 GAARWRHLCLVLL-----GTRGRLAGGSGLIPGSVNDSEGTIDCHADRI	49
RESULT 4		
Q8C9Q6	PRELIMINARY; PRT; 992 AA.	
AC	Q8C9Q4;	
DT	01-MAR-2003 (TREMBlre. 23, Created)	
DT	01-MAR-2003 (TREMBlre. 23, last sequence update)	
DE	Signal peptide	
OS	Kidney muscle (Mouse).	
Qy	62 CQNTPTSYKSCPKCYQGQROBDIDECCNE LNGGVHDLNIPSNYRCTFGJGMLA	120
Db	50 CONAPSKSYKCLCKRGYKGCQCDIDECDNDYNGCVCCTCNPQNYRC-QDFGFLA	109
Qy	121 HOGHNLWDQCLENGCGHTCUNVMGGYECCKEGFFLSDNCNHTCHRSERGUSCNK	160
Db	110 HOGHNLWDQCLENGCGQCIYAMGSYECQHSGFTLSQHNTCHRSNEDANCNK	159
Qy	181 DHGCHICKAPRSVACCRPGFLAKHQDCLTCNHNQGQHSDTAGPPECSE	240
Db	170 DSGCHICRETPKGVAQCDRCPGFDQALQKDCTLTQNYGQCOHSCEBDTGTGCH	229
Qy	241 PQYKATDGSCLREDTVELENTSIVNGDKVRKRLMLTCAVNGGDRCTK	366
Db	230 QKAYPHSDGTCI-----ETCAYNNGGCCTKQ	259
Qy	301 STGMGSCPGFTQZDGKCKD13CQFRNGGDHFCKNIVGFSFDCGGKGKFLLTDER	360
Db	320 TCD:DECFSERTCCHICINSPGFSFOCLHRGTYLYGTHQGYDECSMSNGSDOGCVN	379
Qy	421 TIGSVECOCHGPKHAWKDCD-VKLJLPTSVSPRVSHGSSGGGFCFLR---	474
Db	380 TGSVECVCPGPERLHWNRKDCVEGKLSRAKTSPRAOISCSRAGGVSCLFCPART	439
Qy	475 -----HSGHLS--SDVTTRTSYKLNKGKSL-----	502
Db	440 FVPOSENYVLSGGVPGPQSKALQRNGTSSGJPSCSJADPTPKKARPKRDACKH	499
Qy	503 KNAELFPEGJRPALFEKASVUKESR-----YVNTCSSSKOVCPAAGRSPSTPK	552
Db	500 R----- -----PHSQAKETARQPLDCHVYTFVLCDCSKK--RRRGKSPKE	544
Qy	553 Y-FITVEPELETNQEVTAQDLSCKVTRERKAKRTRKAKRPREPFLGICMNMD	621
Db	545 VSHIIEREFEITKMEASDTCACDCKRKREQLSQAIAKTRKSKIGRQOFQYVNGSTEYE	604
Qy	612 VAKKPRTSBRQAECCGGGAGHAENQCVSCRA GTYDQGGERERCLCPNSTFCNBERGOMTC	671
Db	605 VAKKPRTSBRQAECCGGGAGHAENQCVSCRA GTYDQGGERERCLCPNSTFCNBERGOMTC	663
Qy	672 EPCPRPGNSGALKTTEAMKNSCGGICQSESEYASDGFACQCLALGTFQPBAGRSTCPC	731
Db	664 TPCP--SSDGLIGCARNVSECGQCGCOPFFSAGKPKCACPVGTYQPERGRCFC	720
Qy	732 GCGSATRQGATSFQCETRQCPGCPHFKNTTHRRCRCPVGTQPEFGKRNQVCPNT	791
Db	721 GGSLTLKGEGITSFQCEAKVHFCSSHNM-----HRCRCPVGTQPEFGKRNQVCPNT	780
Qy	792 TTDFOGSTNTQCKONRCGEBLGDTGTYESPNTPGNYPANTECTWTINPPKRIILV	851
Db	781 STDFDOSTNTVHCKQHCGEBLGDTGTYESPNTPGNYPANTECTWTINPPKRIILV	840
Qy	852 PEIFUPIEDGGDYLWKRKISSNSVNTVETCOTYERPIAFTSRSLW	911
Db	841 PEIFUPIEDGGDYLWKRKISSNSVNTVETCOTYERPIAFTSRSLW	900
Qy	912 RQFOVQVYDQDYEGLIEVIRDLYASENHQDILKOKKZKALFDVIAHPOVFKY	971
Db	901 KGFOVQVYDQDYEGLIEVIRDLYASENHQDILKOKKZKALFDVIAHPOVFKY	960
Qy	972 AQESMEMPRSPTRLRSKVSFRFLRKY	999
Db	961 AQESMEMPRSPTRLRSKVSFRFLRKY	988

QY	441 DCVEVKGLPLTSVSPRSVLHCKGSQGGDGFLR-----	DR InterPro; IPR00859; CUB domain.
Rb	442 DCTEPILKCQGSPGASKAMLSNRSKKDTCACTPSRARFLPESSENGFTVSCTPSRRA 455	DR InterPro; IPR06209; EGF like.
Qy	475 ---HSGIHLSSD----VTTTTSVTPKLNKGK--SLKNEAELFEGLRRLPAPEKHS 521	DR InterPro; IPR01491; Thrombomodulin.
Db	456 PARAGHNGNISTNSHICHEAAVLSIKORASPKIKOAKCRUHLRKGTBAGRITGPGAP 515	DR Pfam; PF00431; CUB; 1.
Qy	522 SVKESFRYVNLTCSSGKQVPGAPRSPTRKEM-ITVEELEJENQKETASCOLSICLVR 58C	DR PRINTS; PRO0507; THROMBOMODULIN.
Db	516 CSEQCQTTEHILCKCSRSKGRARRTPKEKEVLTLESEAETAEETASCCIPCRQR 575	DR SMART; SMCC042; CUB; 1.
Qy	581 TEKRILKAIRTLRKAVRHREOFHQLOSGMNIDVAKKPRTSERQA -- ESCGVQGHAENQ 617	DR PROSITE; PS00013; ASX_HYDROXYL; 6.
Db	576 MERRILGSKLKLKSKINQDRLFLRLAGBLDEYLAKRPGVLAGERAEAPMESCRCPOHAGT 635	DR PROSITE; PS01180; CUB; 1.
Qy	638 CVSRAGTYDDGAEERC--CPNGTQNEQGMQTEPSCPENGNCALKPEAWMRNSECGL 697	DR DR PROSITE; PS01187; EGF; 2; 6.
Db	636 CVSPQFGTYVHGQFQCVCPAGTQFOEREGQSDLCFSANGPL--GATWVTCAGC 692	DR SEQ-1-like domain.
Qy	698 CQPERYSAGFAPARCOLAATGFBAGRSFCPPGGLATKHOATSFDCERFVQSCG 757	DR SQ SEQUENCE 961 AA; 103989 MW; 9C6726C4E8BD33E8 CRC64;
Db	693 CQPERYSAGFVSCFKPCPAGTQFQBAEGRIFPPGGLATKHOATSFDCERFVQSCG 752	DR Query Match 58.3%; Score 3262.5; DB 1; Length 961; Gaps 14; Best Local Similarity 60.9%; Pred. No. 1.3e-11; Mismatches 181; Indels 95; Gaps 14; Matches 584; Conservative 99; Mismatches 181; Indels 95; Gaps 14; SeqID 2 GAMAVRWHLSLLALGARGQVVGSSLGLP-----AVDVECESEGTDCHIDAIQNTF 54
Qy	758 HYPNTITHRCRPGTYQFEGKNCVSPGNNTIDFGSTNTTQCCRRCGEGLGEFT 817	DR
Db	753 HYNNNTSIHRIRCAMSQYQDFRQNCFCSRQGKNTDFOGSTVAQCCRQCGELGET 812	DR
Qy	819 SYIEPSPNYPGNYPANTECTWNPPEKRLIVPEIPIEDCGDYLMWRKISSSV 977	DR
Db	913 GYIESPNYPGNYPAGVECINNINPPPERKRLIVWPEIPIEDCGDYLMWRKISSSV 972	DR
Qy	873 TTYECQTYVERPIATRSKLUWQPKSNEGNSARGFQVYVWDEQYQELIDIVRGR 932	DR
Qy	938 LYASENHQEILKDKKLIKAIFDVIAHPQNYFKYTAQESREXPSPFIRLRSKURELR 997	DR
Db	933 LYASENHQEILKDKKLIKAFFEVLAHQPNQFYKTEKHKEMPLSPFIKURSKUSSLR 931	DR
Qy	998 YK 999	DR
Db	992 YK 993	DR
RESULT 6		
Q9ECC6	PRELIMINARY; PRY; 961 AA.	
AC	Q9ECC6;	
DT	2001-01-01 (TREMBrel; 16, Created)	
DT	2001-01-01 (TREMBrel; 16, Last sequence update)	
DT	2003-01-01 (TREMBrel; 23, Last annotation update)	
DE	EGF-related protein SCUB1.	
GN	SCUB1.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mts.	
OX	NCBI_TaxId=10990;	
RN	SEQUENCE FROM N.A.	
R2	STRAIN=JH1;	
RX	MDLINE=0541712; PubMed=11097664;	
RA	Grimmond S., Larder R., Van Hateren N., Siggers P., Hulsebos M.J.M., Arkell R., Greenfield A.; Expression Analysis of a Gene Encoding a Novel Cloning, Mapping, and Expression Analysis of a Gene Encoding a Novel	
RT	Yamalicia: EGF-Related Protein (SCUB1)." ; ! SIMILARITY: CONTAINS 1 CUB DOMAIN.	
RL	Genomics 73:74-81(2000). ; ! EMBL: AF276425; AAC25939.1; -.	
DR	HSSP, P0742; IHCg, DR; MGII:189066; Scubel.	
DR	Inter-Pro; IPR000152; Asx_hydroxyl.	
Qy	564 NQKEVNTASCLSCIVRTEKIKARTTVAHVREOFHQLOSGMNLDAVAKPPT-SERC 623	
Db	557 KYDASSTCEADCMRKRAEQSQQAAKIRKSTGRNQFYVQVLTVEVRAQKALEGT 616	
Qy	624 AESCGVGQGHAENQCVSCTSGRAGTYDGRERILCPNGTFCQNGTCEPFRPGNSAL 683	
Db	617 G-TGIGIQIQLQDGKCYC2CAPGIVYFSDPGCQYFCMEGOLSCCP--SSEG; 672	
Qy	684 KTPEAKWYSECSLCCQPGYESADGAPCQLCAGTQPEAATSCFCGGGLATKHOAT 743	
Db	673 GLAGARNVSECCQQCSPGYFSAQGFRPQACPVGTOPEPGRITGCFCGGGLQKHTGA 732	
Qy	744 SFQDCETRVQSPGHFVNTTHRCIRCPVGTQPEFGKNN-VSCPNTTDFGSTNITO 803	

RA Kawakami B., Suzuki Y., Sugano S., Nagahashi K., Masuno Y., Nagai K.,
 RA Isogai T.; "NEDO human cDNA sequencing project";
 RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RL
 CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 DR EMBL: AK032010; BAC03789 1; -;
 DR InterPro: IPR000859; CUB_domain.
 DR Pfam: PF00431; CUB; 1.
 DR SMART: SM0042; CUB; 1.
 DR PROSITE: PS01180; CUB; 1.
 KW Hypothetical protein.
 SQ SEQUENCE FROM N.A.
 581 AA: 64343 MW: D79C28DAE4EE1 CRC64;
 Query Match: 31.0%; Score 1738.5; DB 4; Length 581;
 Best Local Similarity 55.7%; Pred. No. 4.8e-146;
 Matches 325; Conservative 84; Mismatches 123; Indels 47; Gaps 7;
 QY 459 LHCGKSGGGQGCFIRC-.....HSGLIHS... 433
 2 LSCNRSGKGKOTCACTPSRARFLPESENGTVSCGTPSPRAAPRAGHNCNSTNSHC
 Db 484 -VVTIRTSVIFKLNFGKC-SLKNEELFBGJRALPEGHSSSKESFRYVNLTCSSKG 539
 62 AAVV...KORASIK-KDACKCBLHLRRNGKTEBAG-TG...3APCC... 12;
 Db 540 VPGAPCRPSTPKME-ITVPEFELTENQKEVITASCCLSLCIWKRTEKLRAIRTRKAVER 598
 122 GKGRRAARTPCKEKENVTRLTLEEAVERAETATASCGCLPLCRQMRMLRKLSMRKSIN 191
 QY 599 EQFHQLQSGMNLQVACKPPRISERQ---ESCGVQGHAENOCNSRACTYDQDARECI 655
 182 DRFLRLILAGLQPYELAHKPLVAGERAEPMESCRPGHQHAGTCKVSCPQGQHYHOTEQV 242
 Db 656 LCPNGTQNBBGQMCIECPCPGPGNSALK-PK-BEAMMNSECGLCPGEYSADGPAPCQKA 715
 242 PCPAGFQOERIQGQSDLCIDCSDAHP...-GANTVTTAGCQCCPQHEDGFPCCPQCP 298
 QY 716 LGTFQBEAGRSTSCPPGGGLATKHQATSFQDCETRVQSPGHEFNNTTHRCGPVGY 775
 Db 299 RGTYQCEAEGRTLCFFGGGLT...KHEGAISRFQDCDKVQCPGHEYNTSHRCRAGMY 358
 QY 776 QPFEGKRNQCVSCPGNNTTDFGSTNTIOTKNRRCGGBLGEFTGY-ESPKNPGNFTANTC 835
 359 QPDFRQNCFCSCPQNTSTDFGQSTISVAQCNKRCQCGEBLGFITGY-ESPKNPGNTPAVGC 418
 QY 836 TWTTIPPPKKRILVWPEIPIEDPGDYLVMRKTSSSUTVTTCTYPERIAFTR 835
 Db 419 IANNNPPKKRILVWPEIPIEDPGDYLVMRKTSSSUTVTTCTYPERIAFTR 478
 Qy 896 SKKK...QFKSNNEGNSARGFOQPVYTDQELIEDIVRGRLVASENHOEIKDKKKLTK 955
 Db 479 SRKLUW-INFKUZEANSARGFOQPVYTDQELIEDIVRGRLVASENHOEIKDKKKLTK 538
 QY 956 ALFDVLAHPQNYFKYTAQSEREMFPFSFIRLRSKVRSLRPFYK 999
 Db 539 AFFEVLAHPQNYFKYI-EKHKEMPLPKSFIKURSKVSSFLRPYK 581
 RESULT 9
 Q9TVQ2 PRELIMINARY; PR: 1664 AA.
 AC Q9TVQ2;
 DT 01-MAY-2000 (REMBrel. 13, Created)
 DT 01-MAR-2003 (REMBrel. 23, Last annotation update)
 DE Y64G-0A.7. protein.
 GN Y64G-0A.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chromadozoa; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peicederinae; Caenorhabditis.
 OC NCBI_TAXID=6239;
 RN [.]
 RP SEQUENCE FROM N.A.

RA Mortimore B.J.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN 12;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99059613; PubMed=9951916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for investigating biology";
 RT Science 282:2012-2015(1998).
 RN .3;
 RP SEQUENCE FROM N.A.
 RA Airsccht R;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CR EMBL; Ali11206; CAB60454.1; -;
 DR EMBL; Ali10498; CAB60454.1; JOINED.
 EMBL; Ali10498; CAB57911.1; -;
 DR EMBL; Ali17306; CAB57911.1; JOINED.
 HSP; P0775; IAPQ;
 DR NormPep; Y6G20A.7; CB24549;
 DR InterPro; IPR00019; Asx hydroxyl.
 DR InterPro; IPR00189; EGF_Ca;
 DR InterPro; IPR06239; EGF-like_EG.
 DR InterPro; IPR002349; EGF_CECOVIF_HLICDSR;
 DR PRINTS; PRO031; EGF_LAMIN.
 DR SMRT; SY00179; EGF_Ca; 4.
 DR PROSITE; PS0010; ASX_HYPOXYL; 4.
 DR PROSITE; PS0022; EGF_1; 22.
 DR PROSITE; PS01186; EGF_2; 24.
 DR PROSITE; PS01187; EGF_Ca; 2.
 KW EGF-like domain.
 SQ SEQUENCE 1664 AA; 179275 MW: A69F093B4C705832 CRC64;
 Query Match: 15.6%; Score 875; DB 5; Length 1664;
 Best Local Similarity 25.5%; Pred. No. 1.9e-08;
 Matches 276; Conservative 99; Mismatches 376; Indels 310; Gaps 45;
 QY 55 DCADADLCON---TPTSYK---CSCKPGYQEGEPROD-DIDECONENJNGGVDCDN 104
 Db 78 NCAC-LCHGGTCVPSPSHNDNEQVSCCPVGF-TGKCKQVANECMAN-NGCERACVN 133
 QY 155 ZENYRCICDFERMLADGHNLDVDCI-LENGGCQTCVWNGSYZCCCEGFFEVDNQ 164
 Db 134 TIGTYCRCWPGFELSGJNT-CSDIOCAVNSNGGSDRCVSPGGFRCDTSLYHADG 193
 QY 165 HTC-----IHRSEBGLSMKMDHGCSHICKEAKRPGS-WACECRGFEIAKNGRD- 213
 Db 154 RCQSGFEPENIILKVKTSCTGSDCEHENDENGFEYCRCRVGFKSENKSCOP 253
 QY 214 ILTCHNGNGCCSCHSCDADGPBCSCHQYKHTDQGSCLAREDFTVLETVTSNTTSVGD 273
 Db 254 VDRCFENKGCGQHCTNNHGQAQCQCPFGFLHSYDRRSQVDECAKNGCEHFCENKG 313
 QY 274 DKEVKKR-----LLMETCAVNRNSGCDRCKDTSTGVLH-CSCPVGFTLQDGK 319
 Db 314 TWRKCREGYQLGRDGRTCCEMLGGCOVGNQGQHDCYDQDQGHHVCKRCRNGYILANDQ 373
 QY 323 TCKDIDEQQTNRGCGCFCRQKVCGSDCGKKGKFKLTDSESCQVDECSDLR-TCDHSC 378
 Db 374 LCIDNBHENNGDDCSQVAGSAGVCOCKRJFKRKCED-SECSNRNGGECQ 433
 QY 379 INHGPFTACACNRYTYGFTH-CGDTNECSTNNGGQQCVNTVGSYECCHPGYKHLW 437
 Db 434 SNQEGGGINCSCPGPSELSEDGHSCHOMEECLININGGAQOLCOKRSRCQCFAGYILAH 493
 QY 438 NKKDCV-----EVKSI-----LPTSVSR-----
 Db 494 DEKSCVAASDASDIFSDIEVSKVPGDSDIVEISSIESPADESPPRLVGRHRVKA 553
 Qy 457 -----VS-HQCKSGGPGCFLRGHGIHLLSDV-TIRTWTKEKGKSKLKNELF- 508
 Db 554 CNPQGTSIELFSESENTDPSKCPGPGF---TCQGSCDONGKKSMSRSGLLS 609

QY 509 ---PEG-----RPALEKHS--VKESEF- 52?

DR PROSITE; PS01195; EGF_2; 23.

DR PROSITE; PS01187; EGF_Ca; 5.

Db KCDCPGSGYGEKECQICRGNGWVCAHKCSCKLDBSTGSCRCEDPEKSCDGCFDGFY 669

KW EGFR-like domain.

SEQUENCE 528 165445 MW; 2B48533D8F776E7 CRC64;

QY 528 ---RYNLTCSGGKQVPGAPGRPSPTRKEMFITEVELETRNQKEVTAASCOLSLCVKRT 582

Db 670 GSQONLKCRMDCENGRCOP-VFGYCTPDGLY---GQSCBPCHFTFG 724

QY 583 KRIRKATRTRK-----AVHREQF-----HJQL 605

Db 715 KNRFFPKCARENNSGCBUTSKCRCKPGYGHKKCKMCSPOLPGACAMKSCPAGIRC 774

QY 606 SGMLUDVAKKPPR-----TSEHQAECCGGQGQHAEQNQ----- 638

Db 775 DPVYGDCTRKCPAGYQGNLCOPCPAGYFGYDCEQKSCADVASPHKSKVCHHVGTCTC 834

QY 639 -----VSRAFGTYIDGABERCLCPNGTFQNE-EQWTCIEP-----CP- 675

Db 835 LPGKTGGLQDQSPORNTGPNCAHTC-SCVNGAKCSDSGSCHCTGFGYATCVCPTG 893

QY 676 RPG-----NSGALKTPEAWNNSECGLCOPGEYSADGFAPCQLALGT 718

Db 894 RFGIDCMQCLCKCQNGAICDTNSGSCCECPGWSKCKCOKACAGTFFGACDKCDD-CADM 952

QY 719 F-OPEAGRITSCFP-----GGGL-----ATKHQATTSFO--DCETRV---- 752

Db 953 HCDPSDGEICCPPKKKGKCDETCDGSLFGAGCKGKICSGONGATCDSVTGSCBRCFWRG 1012

QY 753 -----QCSPGKF-----YNTTHC-----IICPVGYQPERGKNN 783

Db 1013 KKCBRCPGDPGRFGQCNATCDCTTNTDSMWNPFVARDCRHVGECRCPAGWTGD---- 1067

QY 784 C-VSPGNNTIDFPGSTNTQKRRGCGELGFITGIESPNVPGNYPANTBC--TWII 839

Db 1068 CQSCP--LGRRHGSCHRISQCSN--GASCRTVGRGDCPSSGMGENCEBCEPGLWGS 1:2:2

QY 94C N 84C

Db 1123 N 1123

RESULT 1:0

08281 PRELIMINARY; PRT; 1574 AA.

ID 08281; DR 01NOV-1998 (Tremblrel. 08, Created); DT 01-NOV-1998 (Tremblrel. 08, Last sequence update); DT 01-MAR-2003 (Tremblrel. 23, Last annotation update) MEGREL.

OS Rattus norvegicus (Rat). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. CX _TAXID:10116; RN 1; SEQUENCE FROM N.A.

RC STRAIN_SPRAGUE-DAWLEY; TISSUE=Brain;

RX YEDLINE-9836089; PubMed=9693030; RA Nakayama Y., Nakai M., Nagase T., Nomura N., Seki N., Ohara O.; RT Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening.;"

DR Genomics 51:27-34(1998); EMBL:BA324621; -; DR HSPB; P00736; IAPC; DR InterPro; IPI00181; Asx_hydroxyl.

DR InterPro; IPI00181; EGF_Ca.

DR InterPro; IP0006709; EGF_like.

DR InterPro; IP0002039; Laminin_EGF.

DR PROSITE; PRO0021; EGFLAMININ.

DR SMART; SNO119; EGF_Ca; 4.

DR PROSITE; PS00022; EGF_1; 23.

Db 109 PGKSKPSPGBCLSJDVDECASANGC--EGP-----CNYCFCPPGPGQLOGBKTCDVD 166

QY 34 PGRRRAAGCQ--EDVDECAQQGDECHADL-CNTPSYKSCPKPY-QGSGRCDDID 88

Db 226 TQHRCOCRQCYQLOEGRGRVRS--PCEAGNGGCHMICEL-RGLAHCSPHGVLA 280

Db 167 EC-RAHNGCQHNTGTSYLGECFGPFLRHDGR-CLAISSCTLNGGCCHQVQLY 225

QY 148 GSYVCCCKRGGFSSDNQTCIHSSEEGJSCMKNGSHICKEAPRSVACRPGFJA 207

Db 208 KNQDCC--LTCHNGGCOHSDTADGPESCHPOMKMTDUGRSCLEREDIVLETES 265

Db 281 ADKTCETEVYDZCALGLAG-AHGCGINTOSSPKVCHASYLGAQROXIE--MEIVNS 337

QY 266 NTTSVWDGKIKRVRKRLMETCAVNGGDRTKOTSTGWHCCPVGFLQJQCKTKDID 325

Db 378 DC-ANSPCCQACANTPGSYESOFAGYRINT-DGGSCEDVDECAASHGCGCHHSNLAGS 416

Db 338 -CEAGNGGCASHGHTSREPLCTPGRYQVLDDEQKTCID 377

QY 326 ECQTRNGGCDPFCRKNIVSFCCGKGRKLJUTBKSCQDVBDSSLDR-TCDASCINERG 384

Db 437 FCGCAGTRDDEDRGTSLESVWLDGRPFVPRJHIAVRLDELRJLQDDYGR 496

QY 401 -CGD-----TNESI-----NNNG-----CQ 415

Db 497 BAAAEAEJRGEBHTTEKFVLDHSFGHDLSTCDCRNGTFCFGQDGDCPCEGWTGILION 556

QY 416 QVC UNTVG--S-YFQOCHRGYKHNWKDCVERKGLJ-LPTVSISPRVSLH--CGKSG 465

Db 557 ENCPDPTEGKNCSSPCTCG-----NESTCDPVLGACRCPGVS--GAHCEDGCPKF 606

QY 466 3GDGFPLRFHSG3HLSUSCVTTIRTSVTFLNEKGK-SKNAELPPEGURPAJPKHSSVK 524

Db 607 YGKSHRKKKCH-----NRECHRHYGACDQPG-VG----- 638

QY 525 ESRFRYKLTCSGGKQVPGAPRSTPKEMFITVFELETRNQKEVTAASCOLSLCVKRTKR 584

Db 639 --RFCHLACPWAFPGP-----CSEDCLC--- 660

QY 585 URKAERTRKAVHBFQHQLQSGNLDVAKKPERTSERQAEASGQGQHAEQNQ-VSCRA 643

Db 661 -BOSH-----TRSNPKDGSCKAQFGQERQAECS 692

QY 644 GTFYIGARECCL-----CANGIF-ONEQGQMC-- 671

Db 693 GPFEGSGCRHCTCOPGVAQDPSVSBECTRCPGPGYQGEGDCGGCCEPGVTCGNSCSCG 752

QY 672 EPCPR-----PNSGALKTPEAWNNSECGLCOPGEYSADGFAPCQJALG-TFOPE 722

Db 753 ARCHVTEGECIPECKG-----EDGADCPCEWRWLGQGCQETCPACHGASNE 832

QY 723 AGTGTGFPCCGGGLAKHQSATSF---QDCETRVQCSGPHFNTTHRCL----- 769

Db 803 TGTCGCLP-3FVGSRCQDTCASAGWYFGCQIRCACANDGHCDPTGRCSCAPGWTGJSC 860

QY 769 --RCVGVYQPE-----FGKNCQVSPG 789

Db 861 GRACISGHWWPDCHPCNCAGHNCDAASG 931

RESULT 11

Q9VS89	PRELIMINARY;	PRT; 1409 AA.	RL [3] Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
ID			RN
Q9VS89;			SEQUENCE FROM N.A.
AC			RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
DT	01-MAY-2000 (REMBrel: 13, Created)		RA Hradecky P., Huang Y., Kamideker J.S., Prochnik S.E., Smith C.D..
DT	01-OCT-2002 (REMBrel: 22, Last sequence update)		RA Tupy J.J., Bergman C., Berman B., Carlson J.W., Celinker S.E., Smith C.D..
DT	01-MAR-2003 (REMBrel: 23, Last annotation update)		RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N., Krommier B., Marshall B., Milburn G., Richter J., Russo S., Seale S.M.J., Smith E., Shu S., Smutnicki F., Whitfield E., Ashburner M., Gebart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
DE	DG7526 protein		RA "Annotation of <i>Drosophila melanogaster</i> genome."
GN			RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
OS	<i>Drosophila melanogaster</i> (Fruit fly).		RN
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oxyphydroidea; Drosophilidae; <i>Drosophila</i> .		RN
OX	NCBI_TAXID=7227;		RN
RN	SEQUENCE FROM N.A.		RN
RC	STRAIN=berkeley;		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.D., Celunker S.E., Holt R.A., Evans C.A., Gocayne J.D., Ammannides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.C., Zhang Q., Chen L.X., Brandon R.C., Rogers R.G., Bieleski M., Chapman M., Pfeiffer B.D., Wang K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., Abtji J.F., Agbayani A., An H.-J., Andrews P.P., Ptakocik C., Baldwin D., Balliew R.M., Basu A., Baxendale J., Andreou P., Baetzke E., Beeson K.Y., Benos P.V., Bernm B.P., Bhandari D., Bolstrakov S., Borokova D., Botchan M.R., Bouck J., Brokskin P., Brottier P., Buttis K.C., Butcher D., Cadieux E., Cerner A., Chandra I., Cherry J.M., Cawley S., Dahake C., Davenport L.B., Davies P., de Pablo J., Delcher A., Deng Z., Mays A.J., Dew I., Dietz S.M., Doston K., Douc L.E., Downes M., Dugan-Rochas S., Dunkov B.C., Dunn P., Dubin K.J., Evangelista C.C., Fultz C., Rocha S., Fleischmann M., Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Gadek A., Geng F., Goracci J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ikegawa C., Jalali M., Kalush F., Karpen G.H., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuljic J., Lai Z., Lasko P., Lei Y., Levinsky A.R., Li J., Li Z., Liang Y., Lin X., Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D., Mekulov G., Milashina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy L., Nutzny D.M., Nelson D.L., Nelson C.R., Nelson K., Nixon K., Nuskeen D.R., Paclob J.M., Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese W.G., Reinert K., Remington K.J., Sanders R.D.C., Schueler F., Shen H., Sme B.C., Szedl-Klamas I., Simpson M., Skupski M.P., Smith T., Spier E., Spalding A.C., Stapleton M., Strong R., Sun S., Svartkas R., Tector C., Turner R., Venter E., Wang K.-H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.-A., Ye J., Yeh R.-F., Zaveri J.S., Zhou M., Zhang G., Zhao Q., Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RT The genome sequence of <i>Drosophila melanogaster</i> ; RL SCIENCE 287: 2185-2195 (2000).; RN [2]		RN
RP	SEQUENCE FROM N.A.		RN
RN	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[15]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter (MAR-2000) to the EMBL/GenBank/DBJ databases.		RN
RN	[14]		RN
RC	SEQUENCE FROM N.A.		RN
RX	YEDLINE-2019606; PubMed=10731132;		RN
RA	Adams M.J., Celunker S.E., Gibbs R.A., Rubin G.M., Venter C.J., Submitter		

CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Rattus.
RN	NCBI_TAXID-116;
RP	Morphogenetic embryonic lung; Dev. Biol. 212:229-242(1999);
RX	Medline:9930231; PubMed:10419698; Email: AFI3506; AAC34439.1; P3555; LENN.
RA	Kanwar Y.S., Cota K., Tian Y., Kumar A., Wada J., Kashihara N., Wallner B., InterPro: IPR00086; Aldehyde_dehydr.
RT	"Cloning of rat fibrillin-2 cDNA and its role in branching morphogenesis of embryonic lung"; InterPro: IPR00152; ASX_hydroxyL; InterPro: IPR00881; EGF_Ca; InterPro: IPR0049; EGF_HL; InterPro: IPR00209; EGF_like; InterPro: IPR00212; Fibril_assoc.
DR	Pfam: PF0008; EGF; 46. Pfam: PF00083; TB; 9. DR SMART; SMC17; EGF_Ca; 42. PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1. PROSITE; PS00022; EGF_1; 2. PROSITE; PS01186; EGF_2; 36. PROSITE; PS01187; EGF_Ca; 43. EGF-like domain. SEQUENCE 2906 AA; 313371 MW; 98664E727044EF58 CRC64;
Query Match	12.8%; Score 720; DB 1; Length 2906; Best Local Similarity 26.9%; Pred. No. 2.7e-54; Matches 256; Conservative 96; Mismatches 308; Indels 292; Gaps 54; Qy 45 DVDECAQGDDCHADALCQNTPSYKSCPKSYOG-EGRCEDDIDECGNELNGGQHDC 102 Db 1192 DNBCGLS-DNLCRNGSKVNMGTYQCSNPQFQATPDRQGSDIDEC-KMNGGSDTC 1249 Qy 153 UNIPGNRRTCDFGFMALHDGHNCLOVDECLNN----GQCQH-CVNMGSYECKEG 157 Db 1250 TNSEGSEYCSCSGEYALMPDGRCADIDEC-BNNPDPIDCGG---QCTNIRPEBYLCYDG 1305 Qy 214 -ILTCRHGNGGC-QHSDDTAJGPECSCHPOYKHMHDGRSCLERDTILEVENTTS 269 Db 1358 TDVDECETGAHNCDMHASCLNPGFKSCREGWV-GNGKCIDDEC-----ANGT. 1408 Db 270 VWDGDKRKRRLMETAVNNGGDR-TKDSTGVMSCPVGFTLQDGKNCQDIDECQT 329 Db 1409 -----HQESIN-----ASCVNTPGSYRACSCBQFT-GDGFTCSVDCAE 1447 Qy 330 RNAGCDH-FCKIVGSDGCCGKFFKULTDEKSCQDVEBCSLDRTDH-SCINHGFFAC 387 Qy 1448 NTNLCEENGQCLNPGAVRCCEMGFPASDRSCQDIDECSFQNMICVFGTNLPGMFC 15C7 Qy 388 AGNRGYTL-YGTHCGTNECS--IN--NGGCOVQNTVSYECOCHPGYKLANKKKDC 442 Db 1508 ICDGFGYGLDRGHCIDCDEADPINCUNG---LQNTPGYRECKCPDPOLFNGTGVC 1563 Qy 443 VEVK-GIPLTSVSPRS-LHCKGSGGGPG--CFLRGHSGHLSSDVFTTTSVTKLNEG 499 Db 1564 VDKRVGNCYLKGFR-----GGSLSCKTEGVGVNSCSCC----- 1599 Qy 499 KSLKIAELFPGLRPALEPKHSSVWESFRYVNLTTSSKLYVGPGRSPKFEVITVE 558 Qy 1600 -CSJGKAWGNPCTCPV-----NSTEVYI-----CPGGG-FRPNQITIE 1641 Qy 559 -----FELTNG-----KVTASD----- 573 Db 1642 DIDECLQLPGLGQGGGNGINTFGSFOCPCQGYLSETRICDIDECFAHPGVCRRGGTCY 1701
DR	DR 574 -----LSCTVRRERKJKAIRDLKAVREOFLQSGEN-LDVAKR----- 615 DR 1702 NTLMNYTCICPPG-----IMQVGGHNCNDMRSKFCVRSYNTTC 1741 DR 616 ---PPRSERQESCGVGQGHENO-CVSC-RAGTY----- 646 DR 647 YDGRERCLCPQTFQNEGOYTCERCPRPQNGALKTPAWN-----NSECG-- 695 DR 1801 ICCCCPISGICANVCTQHNGSPRC-CP-----TGFNDLLVUCEDCNGD 1850 DR 696 GLCQ-----PEVSADGAFBOLGLG-TPQEAS--RTSCP---CGGGLATPF 739 DR 1851 NCORNADCGINSPIRE-----GAGFKLSPNGAVDRNCJSPNNSHGLV 1903 DR 740 CGA-SFQ-DCETRVCSSGHTYNT-MRCIR-CPVTCYDPEBKRNKNVSCG-----N 792 DR 1904 CG-SYOCION:KASDCTMCMVDSBCCRHPGNG-CKTNTSYGNCYCPGELTHN 1961 DR 791 TTDPFGSTN-TQCKURRGCGELGF-----TGY-ESPNNPGYNTANTEC 835 DR 1962 DCQDIDCSCSFROVORKRGCNEIGPKLNLNEYE-CPGKNCIDTN-EC 2012
Db	RESULT 24
Qy	096cPa PRELIMINARY; PRT; 2809 AA.
AC	996JP8
DT	01-DEC-2001 (Tremblair. 19. Created; 01-MAR-2003 (Tremblair. 23. Last annotation update; DE Fibrillar3. GN KIAA1776 OS Homo sapiens (Human). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-eostomi; Yamaia; Eutheria; Primates; Catarrhini; Hominidae; Homo. OC _TAXID=9606; RN 1. RP SEQUENCE FROM N.A. RC TISSUE-Brain; MEDLINE:21247-30; PubMed:11347906; RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Charra O.; RT "Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."; RT DNA Res. 8:85-95(2001); RA Res. AB053470; BAB47-08.1; --. DR EMBL: AB053470; BAB47-08.1; --. DR InterPro: IPR00152; ASX_hydroxyL. DR IPR01881; EGF_Ca. DR InterPro: IPR006209; EGF_like. DR InterPro: IPR00212; Fibril_assoc. DR PROSITE; PS0008; EGF; 45. DR PROSITE; PS00683; TB; 9. DR SMART; SMC17; EGF_Ca; 41. DR PROSITE; PS00010; ASX_HYDROXYL; 41. DR PROSITE; PS00022; EGF_1; 2. DR PROSITE; PS00022; EGF_2; 36. DR PROSITE; PS01187; EGF_Ca; 40. DR EGF-like domain. SEQUENCE 2809 AA; 300323 MW; 20204CC006C0161F CRC64;
Query Match	12.7%; Score 710.5; DB 4; Length 2809; Best Local Similarity 24.8%; Pred. No. 1.8e-53; Matches 262; Conservative 79; Mismatches 316; Indels 401; Gaps 50; Qy 45 DVDECAQGDDCHADALCQNTPSYKSCPKSYOG-EGRCEDDIDECG----- 91 Db 1071 DVDSGARDPLCIGR-GTNTNTDGSYKQCPGSHBTAKGTAEDCIDECSLSDGGLCPHGQC 1129 Qy 92 -----NE-----NGG-VHDLNIPNRYCTCDFGFMALH 121

Db 1130 VNVIGAFCSCHAGFQSTPRQGCVDENCRVNQGGDVHCNTESYRCSCQGSVSLMF 1189
 Qy 122 DGHNCLDVCECLENNGGC-QHTCNVMGSYECCKEGRFLSBNQHCTIERSSEGLSMKN 180
 Db 1190 DGRACADDECEENPRVQDQHATNMPPGHRCLYDGFMATDMRMTDVDE--CDLN 1245
 Qy 181 DHGCSHICKEAPRGSSVACECRPGFELAKNQRC--ILTCNHNGGQH- SCDTAGPE 236
 Db 1246 PHCLHGDCENTKGSFVUCHQGKATGCSDVDECEVGGHNCDSHASCNTPGSFS 1305
 Qy 237 CSCHPOQKMTDGSCLSREDTWLETSNTSVWDG2KRVGRLLMETCAVNNGGCDT 296
 Db 1306 CRCPGMWV--GDFECHOELDCECSQEHRCSPR---GD----- 1337
 Qy 297 CKDSTGWHCSCPVGFTLQDGSKTKO1DECOTRNGGDDH-FCKNTVSGDGCCKGFKL 355
 Db 1338 CLNYPGSGRTCTCGFA -GDGFCCEDRDECAENVDLNGGCGCINAPGGYCECEMGFP 1395
 Qy 356 LTDEBSCDVSCESSLDRCDH-SCINHNGTACACNRGTYL-YGFPHGDTNECS--IN- 410
 Db 1395 TECHRACDVPDECHOGNLCAFSGENLPGMFRCTCINGYELDGRGGNCTDNECAVPNC 1455
 Qy 411 -NGCQQQVQUNTVOSYEQCHPKYKLHNKKOCVEVKGLLPTSVSPRSVLHGKSSGGD 469
 Db 1456 ING---VCOINTPSYLCSCPQDPFELNPVGCVETR-----AGN 1491
 Qy 470 CFURCH---SGHLSSD--VTIRTSVTFKUNGEKSLSKNAELFFPSSJRPALPEHSS 523
 Db 1492 CFEIETHDGDGSCSABIGVGUTRASCC----CSIGRAWNPOLCP----M 1535
 Qy 524 KESRYVNLTCSKSKQVQGAPG-RPSPTRKEMP-TVEFE-ETNQKEVTPASCP-LSCIVRT 581
 Db 1537 ANTEYIFL-----CPGEFGCPNR -----IVVIAEDIDSCQELPGICGSDV--- 1586
 Qy 582 EKRKARTPLTKAVHPRQFHLLASGMOLQVAKRR--RTSE---RQAESCGVQGCHA 634
 Db 1581 -----NTFGSFQCCCPGSHLSEHTRCEDZBCTSHG-- 1614
 Qy 635 ENQCVSRA GTYVUGARERCLP-----NGTFQE----- 665
 Db 1615 -----ICGGTCNTLGNVTCVPAEY-QVNGGNNCMDMRKSVCVERHYNGCQNELAENP 1669
 Qy 666 -----EGQ---MTCEPCPRP-----GNSCA-----LKTPERAMWNSCC-- 695
 Db 1670 TRKNCCCSYNIGQWNRPEACEPTPISTDYQIQLIGNQARGF-ZDTH3KPLDCESEIP 1723
 Qy 696 -----GLC---QPGYSAD---GF-----APCC----- 712
 Db 1730 AICANGICNOIGSFRCECPAGFYNS-LLACEDVDECGSRASPCQONADCINPPSIRC 1799
 Qy 713 -----L-----L-----L-----CALGTPCPEA3RTSCF-- 729
 Db 1790 KCTR3YKUSPGGAVGQNCBREIIPVCSHGDCMDESYMCJHRG-FQASADOT-CYDI 1849
 Qy 730 -----PGCG-----L-----LATKGAGTSQDGETRV-QCSRGHFNIT- 763
 Db 1849 DECURQPCGNGTCHNIISYNCNLCPGPFWVTHNGCVDPECTLGVCGFRCHTAG 1909
 Qy 764 -----THFCI---RCPGTVQPEFGKNCVSPGKTTDFGS 798
 Db 1909 SFHICQDFELTDAGKNDVNTNCISLJAGTCFGTCONLEGFSRCTI- CPPSFQVQSCHC 1967
 Qy 799 TNITOCKNR-----RCGELGPF----TGYIESPN 524
 Db 1968 IDIECSEPNLCLFGTCINSPGSSQCLCPPGFLSDLN 2005

RESULT 15

Q8435 PRELIMINARY PRT 810 AA.
 ID Q8435; AC Q8435; DT 01-MAR-2003 (TreMbare, 23, Created)
 OC NCBI_TaxID=0090;
 RN [1] _SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
 RX MEDLINE-2235463; PubMed=1246651;
 RA THE RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RU Nature 420:563-573(2002).
 DR EMBL: AK08351; Bac38785.1; -
 SQ SEQUENCE 810 AA; 90072 MW; 12CB03938C27BC CRC64;

Query Match 12.4%; Score 695; DB 11; Length 910;
 Best Local Similarity 30.0%; Pred. No. 8e-53; Mismatches 224; Indels 78; Gaps 14;
 Matches 160; Conservative 71; Warnings 0; Gaps 0;
 Qy 62 CKRPTSYKSCPKGY- CGEGRGCEDECGNELLNGGCVHCCNINQYRCCFDGFM; 119
 Db 253 CNT-PGSPSYICKKOGYI-STDKKCRIDCATC-DHGCEOLVNMGSFVQCGSYT; 311
 Qy 120 ARDHGNCIDVCECLENNGGCQH-TVNMGSYECCKEGRPGFELAKNQRC--ILTCNHNGGQHSCDAGSPEC 179
 Db 312 AEGKAKT-NVQYCASENHCHECNCVNAESESYLORCHEGFAALSKPSKIC---YCAS 367
 Qy 180 RDHGSCHICKEAPGSVACECRPGFELAKNQRC--ILTCNHNGGQHSCDAGSPEC 217
 Db 368 SNHGCQHCVNA-QTSALECRJCKGFMNLNDRKTRNLYCAJNKPGCBEHCNVTEEGYC 426
 Qy 239 SCHPOQKMTGRCLE-----RECTYLETSNTSVDG--DKRKRRLME 284
 Db 427 RCGCGNLPKPGKTCRSVHDCAQDGHGECOLCUNTEESFVCCGSEGLINDLKTCSRAD 486
 Qy 285 TCAVINGGORTCKDTSVH-CSPVPGTLQDCKTKO1DECQTRNGCDDHCKNIVG 344
 Db 487 YCLLUSNNGGEBVSNTDKSFACCPGEPVHLRSPGKTCAKDAGHGCHECNSSEBD 546
 Qy 345 PGCGRKGRKJLDEKSCQDDEC-SLRTRCHSNCINHGTGACNNGTYL-YGFHFG 402
 Db 547 FVCGCPREGYIIRCGKCRKKEVQCDVNRHGCGHUCVNGSGESYVCKLEGPRKAEDGKCS 606
 Qy 403 CTNEBSLNN3QCVQVNTVSGFCQCHPGYKAKNKKOCVE----- 444
 Db 607 RKNVCKSTQHGEEMCVANGNFSLRCSEGFIJAEDGKCKRCTEGPDJLVNFDGSKL 666
 Qy 445 -----VKGJLPT-SVSRVSHGKSGGDDGFLRCHSGIHHSDDVTRTSM 492
 Db 667 JEENEFETVKHFTGJGSLAVSPKA-----ARVGLQYSQVRTEF 708
 Qy 493 FK---LNECKSLKNAELPPEGRPALPEKHKSVKSFYRIVNLTCSGGKQV 541
 Db 709 JRGESSAKENKAVTHMKYKMGKSMGTMGLKH-MFERSP70VEGARPPS70V 760

Search completed: October 22, 2003, 11:30:17
 Job time : 130 secs